DOCUMENT RESUME

ED 208 094 UD 021 700

AUTHOR	Stickney, Benjamin; Marcus, Larry
TITLE	The Limits of Compensatory Education.
PUB DATE	Apr 81
NOTE	108p.; Paper presented at the Annual Meeting of the American Educational Research Association (Los Angeles, CA, April 13-17, 1981). Not available in paper copy due to authors' restriction.
EDRS PRICE	MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS	Academic Achievement; *Black Achievement; Black Students; *Compensatory Education; Educationally Disadvantaged; *Educational Opportunities; Elementary Secondary Education; *Equal Education; Equal Opportunities (Jobs); Federal Programs; *Program Effectiveness; *Program Evaluation
IDENTIFIERS	Coleman (James S); Elementary Secondary Education Act Title I; Project Head Start

ABSTRACT

An overview of compensatory education programs is presented in this paper. The rationale for compensatory education is outlined, and the early implementation of these projects at the local and national levels is described. Specifically addressed are Project Head Start and Title I programs. The paper also discusses James Coleman's "Equality of Educational Opportunity Survey" and reviews the major studies of compensatory education conducted at both national and local levels from 1965 through 1980. It is proposed that a review of the more scientifically promising evaluations of compensatory education can yield insight into the potential to redress racial differences through academic achievement. Such a review, it is said, might also reveal the most productive means of increasing the effectiveness of the public schools for underachieving black pupils. The paper itself is abstracted from a larger work by the authors titled "Race and Education: The Unending Controversy." (Author/APM)

****
* . Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED208094

THE LIMITS OF COMPENSATORY EDUCATION

Benjamin Stickney

Pikes Peak Board of Cooperative Services

Larry Marcus

New Jersey Department of Higher Education

Paper presented at the American Educational Research Association Annual Meeting, Los Angeles, California, April 13-17th, 1981.

This paper is abstracted from the book Race and Education: The Unending Controversy, by Larry Marcus and Benjamin Stickney (Charles E. Thomas Publishers, August 1981).

UD021700

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.

- Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY

Benjamin Stickney

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

THE LIMITS OF COMPENSATORY EDUCATION

Part I

"TOWARD EQUALITY OF EDUCATIONAL OPPORTUNITY: HEAD START, TITLE I, THE COLEMAN REPORT"

During the decade following the 1954 landmark Brown vs. Board of Education Supreme Court Decision, a growing number of American educators experimented with alternative means of educating poverty stricken underachieving children. The social and educational issues publicized by the desegregation struggle seemed to be the impetus for much greater attention being given to the role of the schools, whether segregated or integrated, in promoting equality of educational opportunity and achievement. Throughout the late 1950s and early 1960s, school districts across the country initiated and expanded what was often labeled "compensatory education" programs for "culturally deprived" children. By the middle sixties the local initiatives had swelled to national commitments, manifested by Congress legislating Head Start for impoverished pre-schoolers in 1964, the Elementary and Secondary Education Act (Title I) for "educationally deprived children" in 1965 and ordering a massive study of educational inequalities (the Coleman Report) also in 1965. This chapter will review the rationale for compensatory education, describe the early implementation of these projects at the local and national levels and review the startling conclusions of perhaps the most controversial and influential eduational study in American history, James Coleman's Equality of Educational Opportunity Survey.

THE RATIONALE FOR COMPENSATORY EDUCATION TO 1965

In reviewing much of the social scientific and educational literature which was used to establish the wisdom of such programs as Head Start, one can conclude that the following three fundamental assumptions constituted a broad rationale for compensatory education: (1) The environment had a profound influence on children's measured intelligence and school achievement, (2) schools constituted an important part of children's environment, and (3) schooling could compensate for the environmentally determined deficiencies in measured intelligence and pupil achievement. In the late 1950s and early 1960s, the process of schooling was often viewed as largely irrelevant for much of the nation's poor and "culturally deprived" school population. Among the black population racism and poverty were typically seen as unusually dehumanizing and as stifling normal intellectual growth and academic progress. Moreover, the schools aided and abetted black underachievement by their adherance to white middle class values, usually manifested by the employment of racially and culturally biased teachers, methods and materials. Therefore, it seemed appropriate to design schooling experiences for millions of "culturally deprived" and educationally disadvantaged children which gave greater recognition to, and fit more comfortably within, the childrens' socio-cultural milieu. For black youth, desirable educational experiences often involved concentrated efforts to improve teacher-pupil rapport. Affective components such as these combined with individualized and novel techniques in the cognitive areas would compensate greatly for environmental inadequacies and thereby maximize the chances of accelerating cognition and achievement.

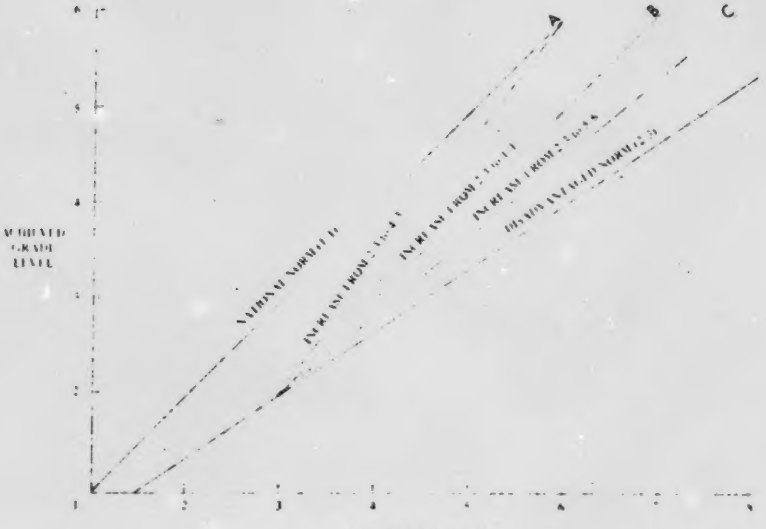
Cumulative Deficit

In the late 1950s and 1960s, the longer most poverty stricken children attended school the greater was their need for accelerated learning to reach the average scholastic levels nationally. It had been reported that the I.Q.s of children from low income families typically declined throughout their years of schooling, and the term "cumulative deficit" emerged as a commonly employed label for this hypothesis.[1] On the various standardized reading an[illegible]th tests, most poor school children would fall further and further behind the national average, a phenomenon which was not necessarily caused by any absolute decline in cognitive capacities but by the cumulative effects of different learning rates.

The cumulative deficit in achievement was perhaps most simply illustrated by David Hawkridge and Michael Wargo and their staff at the American Institute for Research (AIR), who published a series of reports from 1968-1972 for the U.S. Office of Education (OE) on exemplary compensatory education programs (see Figure 1).[2] According to these AIR researchers, if one used achievement per month of instruction as the measurement criteria, it would be discovered that the achievement rate of "badly disadvantaged" children, as measured by standardized reading and math tests, would be approximately two-thirds (2:3) that of average children who learned at a 1:1 rate.[3] At this growth rate in reading and math, one could expect the average "badly disadvantaged" child to have fallen approximately one year behind the national norm at the end of the third grade, two years behind at the end of the sixth grade, three years behind after the ninth and four years below grade level if the student graduated high school. (See Figure 1

Since the AIR studies, there has been considerable criticism of using achievement per month of instruction and grade equivalencies as the longitudinal measurement criteria,[4] and a growing controversy over how much of the cumulative deficit can be explained by possible differences in social class-related learning rates over the summer (see Chapter 5). There has also been data which challenges the two-thirds achievement rate generalization for such a nebulous description as "badly disadvantaged" children.[5] But whatever the technical flaws of the AIR analysis, the relevant data from the 1950s and 1960s were virtually unanimous in suggesting that children from low income families in general, and impoverished black children in particular, were typically slower in acquiring the basic skills than their middle class counterparts and that the resulting deficit was indeed cumulative.

Consequently, in the late fifties and early sixties, a growing number of educators became concerned with both preventing and reducing the scholastic retardation of "culturally deprived" children. At the pre-school level, compensatory education programs often attempted to raise the I.Q.s of lower class children in order to



Legends for Marcus and Stickney – RACE AND EDUCATION
Figure 1. Learning rates children achieving at the national norm and the disadvantaged norm and three possible measures of program impact. Early compensatory educator evaluators tended to favor impact A in assessing program effectiveness. Figure adopted from M. Wargo et al.: *Further Examination of Exemplary Programs for Educating Disadvantaged Children.* Arlington, VA: ERIC Document Reproduction Service, ED 055 128, 1971, p. 15.

improve their chances of normal achievement growth after entering elementary school. At the elementary and secondary school levels, compensatory programs concentrated more on accelerating the achievement of disadvantaged children than on raising the I.Q. Therefore, one cognitive objective of the pre-school programs was to prevent the cumulative achievement deficit from ever occurring, while the programs for older children proposed to reduce or even eliminate the progressive achievement gap. The early U.S. Office of Education AIR investigations noted that for disadvantaged children who had already fallen behind, the achievement deficit could only be eliminated if special programs raised the rate of gain of these children to a point that exceeded the national growth rate of a month of learning per month of instruction (1:1).

> "At a month-for-month rate, they will neither catch up nor fall farther behind but . . . rate of gain must be greater than month-for-month if disadvantaged children are ever to reach the point where they can perform at grade level (e.g., . . . increase from 2:3 to 4:3). Since this is the ultimate objective of compensatory education, an achievement rate of gain greater than month-for-month was established as the criterion for educational significance."
> (See Graph One)

According to AIR, "educational significance" meant program success and how successful compensatory education programs have been in significantly reducing the class related inequalities in academic achievement is an issue we will address in Chapter Five. For now, it is important to continue reviewing the rationale for the many enrichment programs by taking a closer look at the underlying assumptions regarding the relationship among the environment, the schools and learning.

The Environment and Cognition

At the time large scale compensatory education efforts were legislated at the federal level, there existed a preponderance of opinion in the literature contending that the environment had considerable influence on measured intelligence and school achievement. Interpretations of the data by many social scientists expressed the interactionist position on the development of human intelligence, which assumed that a child's cognition was the product of complex encounters between his genetic endowment and the environment's unfolding of these innate abilities.

Two of the more influential reviews of the literature were published in book form as *Intelligence and Experience* by J. McVicker Hunt[6] and *Stability and Change in Human Characteristics* by Benjamin Bloom.[7] Between these two volumes, educators could address the nature-nurture question by drawing on studies of animals, twins, siblings, foster children, institutionalized children and similar populations over time.

If one reads the studies reviewed by Hunt and Bloom with any objectivity, it is difficult to avoid emerging with the strong impression that heredity plays a major role in explaining the variation in measured intelligence among persons in a given population. There were simply too many cases of identical twins, who had been separated in early childhood, scoring strikingly similar on I.Q. tests when reunited as adults[8] and of unrelated children reared together, having lower intelligence correlations than siblings reared apart.[9] One of the most frequently cited twin studies was conducted in the 1930s by Horatio Newman, Frank Freeman and Karl Holzinger (the Chicago Group), all from the University of Chicago. These researchers collected longitudinal intelligence data on nineteen pairs* of identical twins separated in most cases before the age of two.[10] When tested in adulthood, the average I.Q. differnece of thirteen pairs raised in rather similar environments was only 4.4 points which is approximately the same as the 2 to 3 point average I.Q. difference of identical twins reared together.[11] In reviewing studies of two groups of siblings reared apart and five groups of unrelated children reared together, one finds more position intelligence test score correlations for the separated siblings (typically falling in the +.30 to +.50 range) than for the commonly reared unrelated children (averaging somewhere between +.10 and +.30).[12]

If the research data made a strong case for the influence of heredity on determining measured intelligence differences, how could advocates of compensatory education emerge from their reading of Hunt and Bloom with an environmentalist interpretation? The answer seemed to lie in the degree of the environmental differences. For example, if two identical twins were separated as toddlers and raised in rather similar environments the I.Q. differences at adulthood rarely exceeded 7 points.[13] When comparing this figure to the roughly 12 point I.Q. difference between fraternal twins reared together and among siblings raised in the same home, and the 15-16 point difference found between genetically unrelated children reared together,[14] it would appear, at least from this analysis, that environmental differences within a given community account for a relatively small fraction of the average I.Q. differences which exist among persons of that community. However, if the environmental differences are substantial, there is evidence that persons with identical genetic endowment will differ considerably in measured intelligence. A closer look at the Chicago Group twin data will reveal that six of the twenty pairs reared in the most contrasting social environments averaged 12.5 I.Q. points difference and the six differing greatest in education differed by an average of 15 I.Q. points.[15]

---

*The Chicago Group actually reported twenty pairs, but because number 20 did not include data on social and physical environmental advantages, it is excluded from analysis in this volume.

Both Hunt and Bloom use the correlations between the contrasting socioeducational environments and substantial I.Q. differences reported in the Chicago Group study as providing evidence to support the plasticity of intelligence.[16] They also refer to reports of children raised in institutions as making a strong case for environmental influence. One of the more prominent studies was conducted by Rene Spitz, who in the middle 1940s measured the effects of maternal deprivation and "hospitalism" on the psychomotor, affective and cognitive behavior of children during their first year of life, reared in two contrasting Latin American institutions. Documenting behavior by films and by performances on the Hertzer-Wolf baby tests, Spitz concluded that the considerable maternal contact provided by one institution (a "nursery" attached to a reformatory for delinquent young women) was largely responsible for what appeared to be normal infant development while the relatively limited maternal and other adult contacts of a "foundling home" for children of economically destitute mothers explained the declining developmental quotient and increasing susceptibility to diseases of these poorly stimulated infants.[17] The Spitz observations gained some support from the work of Ruth Levy who in 1947 reported on comparisons of orphans raised in boarding homes with similar children reared in nurseries. At the age of two and one-half years, the boarding children, who received the greater adult attention, were clearly superior to the nursery children on a "developmental quotient" index.[18]

By the mid-1960s perhaps the most convincing evidence supporting the contention that intelligence is not "fixed" had come from the reports of Harold Skeels, who studied not only the effects of institutionalism in early childhood but also its longitudinal influence on behavior. In the original study, thirteen of twenty-five "mentally retarded" orphans residing in an "affectionless" institution were removed by the age of 18 months to another orphanage where they had intimate contact with several mildly mentally retarded women. The remaining twelve children, constituting a contrast group, remained in the original orphanage. After periods ranging from several months to over two years (depending upon the time of removal and testing) the mean I.Q. of the experimental group improved dramatically (64.3 to 91.8) while the measured intelligence of the contrast group suffered considerable decline (86.7 to 60.5).[19]

Because of the low reliability of intelligence tests at such an early age, methodological flaws and controversy over statistical regression, the Skeels work (coauthored with H.M. Dye in 1939) was less than convincing until more recent reports in 1965 and 1966 revealed the results of follow-up interviews with all the subjects some twenty years later. Eleven of the thirteen children in the experimental group were adopted, and all were self-supporting leading normal productive lives, virtually indistinguishable from the general population. On the other hand, subjects in the contrast group either remained wards of institutions or were generally living marginal lives on the outside as unemployed dependents or unskilled laborers.[20]

In comparing the educational attainments of the two groups, Skeels reported that "In the experimental group the median grade completed was the twelfth; in the contrast group, the third" and that one of four experimental group subjects attending college for one or more years received a B.A. degree. One of the adopted girls ("who initially had an I.Q. of 35 . . ." and might ". . . have been sterilized in late adolescence or early adulthood" if she had been in the contrast group) had graduated high school, attended college for a term and was married with two sons, who had I.Q.s of 128 and 107.[21] Unfortunately, Skeels did not administer I.Q. tests in adulthood to any of the subjects.

By the mid-1960s the large scale longitudinal studies of intelligence, which included I.Q. test data collected over a number of years, were in general agreement that the I.Q. of most persons varied very little after about four years of age.[22] In his influential publication Stability and Change in Human Characteristics, Benjamin Bloom reviewed the major multi-year studies to date and concluded that ". . .intelligence develops in an exceedingly lawful way. . .," which involves a ". . .rapid increase in the correlation between the criterion measure and measurements made in the early years and a less rapid increase after age four."[23] According to Bloom, the correlations between I.Q. at age 17 and at age 2 is +.41, at age 4 it is +.71 and by age 11 is +.92.[24] In other words, after the critical period of early childhood, measured intelligence becomes a relatively stable characteristic, at least until maturity. Bloom's analysis of an infant's susceptibility to environmental influence will be addressed later in this chapter.

Since most subjects in the major longitudinal studies were middle income and probably grew to maturity in environments which were rather stable, it may come as no surprise that the I.Q. scores varied little between the ages four and seventeen. (In reviewing the major longitudinal studies in Intelligence and Experience, Hunt cites two cases--one from the California guidance study and another from the Berkeley study--in which poor health and maternal overindulgence were thought responsible for respective declines in I.Q. from 140 to 85 between ages 2 and 10 and from 133 to 77 between 2 and 18).[25] But what of studies of persons whose environment has undergone considerable changes over the years or reports of children reared in atypical surroundings? Bloom and Hunt both cite the research of Otto Klineberg which concluded that the length of residence by blacks in New York City, who had migrated to that point from the South in the 1920s and early 1930s, correlated positively with the growth of measured intelligence.[26] They also included a study published by E.S. Lee in 1951, which reported a significant increase in the I.Q. of Southern black children after they moved to Philadelphia.[27]

Most research on poverty stricken and black children, however, reported a gradual decline in children's I.Q.s as long as they resided in a culture of poverty. There was a study reporting a 20 point I.Q. decline (100-80) of Appalachian children between ages six and sixteen[28] and a number of studies showing a similar cognitive decline in samples of the black population. Therefore, it was the environmentalist position that the moderate correlation (commonly in the area of +.25 to +.50) between measured intelligence and social class could be largely attributed to the depressing effect of certain deprivations in the cultures of poverty.

But what was viewed as a culture of poverty and what were the environmental variables within it which may have contributed to the stifling of intellectual growth? Perhaps the most interesting and influential analysis of the phenomenon in the late 1950s and early to middle 1960s was made by anthropologist Oscar Lewis. Basing his generalizations on a number of cross-cultural observations (most extensively conducted in Mexico, Puerto Rico and New York City), Lewis contended that a "sub-culture of poverty" existed in many places in the world, particularly in "class stratified, highly individuated, capitalistic" societies. Aware of middle class values and cognizant of the difficulty of upward mobility, persons sharing Lewis' sub-culture of poverty had feelings of hopelessness, had negative self-images and suspicious attitudes toward such dominant culture controlled institutions as the police and the schools. Such alienation, despair and economic want existing aside relative affluence, contributed a number of poverty culture characteristics, most prominent of which included matricentrism, present-time orientation, impulsiveness, irrationality, disorganization and general authoritarianism. Since these cultural variables were self-perpetuating and therefore contributed to the maintenance of ignorance and destitution, Lewis concluded that ". . . the poverty of culture is one of the crucial aspects of the culture of poverty"[29] and that it would be best "transformed and eliminated."[30]

Despite Lewis' contention in 1965 that the culture of poverty had been largely eliminated in the United States (including only six to ten million people) and a number of reports which took issue with Lewis on such suggested poverty-related variables as "feelings of hopelessness," the flavor of his writing typifies a number of other important manuscripts of this period such as The Culturally Deprived Child by Frank Riessman (1961),[31] Slums and Suburbs by J.B. Conant (1959)[32] and The Dark Ghetto by Kenneth Clark (1965).[33] Theoretically, poverty was viewed as intricately interwoven with many attitudes and behavioral patterns which often severely impaired a child's chances to acquire those aptitudes and skills which permit one to succeed in school. Lewis' anthropological studies seemed to have given a certain credibility to this "deficit model" (which was often interpreted as putting the blame for poverty and poor school achievement on poor people themselves) and could be interpreted as providing a tidy explanation for such platitudes that the poor (particularly the black poor) were relatively deprived of motivation, of language and/or learning stimuli.

In the affective area, the dehumanizing conditions permeating the black experience in the United States was typically viewed as contributing to low self-concept and aspirations.[34] Without strong egos black children could not make maximum use of their intellectual abilities and without a promising view of the future, black youth could hardly be expected to care much about occupational status or getting ahead someday. These feelings left black school children very poorly motivated, thus contributing to their underachievement.

As plausible as this argument may sound, the social scientific research did not necessarily support the contentions that blacks had weak egos or low aspirations. In 1963 Ausubel and Ausubel reviewed the research on ego development of black children and argued emphatically that the majority of segregated black children in American society, perceived themselves as "an object of derision and disparagement." The authors based their conclusion mainly on the perceptions of several observers and a handful of studies of children's color preferences.[35]

But the material they cite provides far from convincing evidence. For example, one of the most frequently mentioned series of studies were those conducted by Kenneth and Mamie Clark in the 1940s on the doll color preference of Northern and Southern Negro children (see Chapter One). The Clark's work had unexpectedly found that Northern Negroes were more likely to accept a "white" doll than their Southern counterparts. The Clarks had argued that the difference may have been caused by ". . . the fact that in this sample there are many more light colored children in the North than there are in the South."[36] However, it could have been just as easily argued that the Negro children in the more racially integrated Northern city (Springfield, Massachusetts) may have had greater interaction with white persons than the children in the Southern sample, which served to broaden the former group's frame of reference in terms of social status beyond the immediate black community. Although they may have been more integrated with whites, there was little question which race in Springfield in the 1940s enjoyed the greater social status. On the other hand, in a highly segregated Southern community, a black child's frame of social reference may have only included his immediate racial group, and his self-concept may have depended largely upon his relative position within that social entity. A study conducted in the early 1960s of black and white attendants in a mental hospital may serve to illustrate this point. On a self-esteem scale, it was reported that the black employees had even a higher self-image than the white workers.

> "In this middle Atlantic city, the job of attendant is a relatively good job for a Negro but a very poor position for a white. Self-esteem may be more a matter of one's position within one group than the rank of the group in relation to other groups." 37

In short, by the middle 1960s there existed little, if any, solid evidence that black children had a lower general self-concept than white children, and even if they did, the research data showed only that self-concept correlated positively with cognitive achievement. It could be argued just as easily that higher achievement caused higher self-concept as it could that higher self-concept contributed to greater achievement.

There was also research to challenge the assumption that black children had lower aspirations[38] and indeed the Ausubels' review had concluded that pencil and paper measures indicated very similar motivational levels of black and white children. In their commentary on the motivational findings, the Ausubels probably echoed the feelings of many social scientists when they dismissed the results as invalid because of the "unrealistic" defensiveness of black children. While this analysis might have been plausible, it was at best speculative. Accordingly, compensatory education programs that used the findings of empirical inquiry as justification for their emphasizing the raising of self-concept and aspiration levels to improve motivation instead of prioritizing the actual teaching of reading were probably going well beyond the data.

Regarding the language of poverty, few could argue with the notion that a child whose native language or dialect was different from that used in the school usually faced a greater handicap than his counterpart who was reared speaking standard American English. Many educators went a step further, however, by suggesting that the speech of the impoverished actually inhibited cognition.[39] Probably the most frequently cited research was that conducted by British sociologist Basil Bernstein who in 1962 labeled as "restrictive" the linguistic code of many working class families and "elaborated" the dialect more often associated with middle class families. According to Bernstein, characteristics of a restricted language include short or incomplete sentences, a repetitive use of conjunctions, and a limited use of subordinate clauses, modifiers and the impersonal pronoun. Described as a language of implicit rather than explicit meaning, the restricted code may retard the ability to think abstractly.[40]

In this country several prominent educators reached linguistic conclusions similar in tone to those of Bernstein. In 1964, David Ausubel viewed complex conceptualization as more difficult for typical lower class children because of their constant shifting concrete to abstract modes of thought and comprehension.[41] A year earlier Vera John concluded that the language of middle class black children was more adult-like than the language of lower class black youth. Because of greater middle class parent-child verbal interaction during childhood, John felt that "the middle class child had an advantage over lower class in tasks requiring precise and abstract language."[42]

Despite condemnation by many linguists and a later denial by Bernstein that his work had any relevance for compensatory education, many educators were convinced that poor children (particularly black poor children) needed some sort of systematic language training.[43] Martin Deutsch sounded a popular theme when he stated in 1964 that lower class pre-schoolers needed a "language training program where words are repeatedly placed in meaningful context. . ." Drawing on Bernstein and his own research in Harlem, Deutsch noted that lower class children are not restricted in expressive language ability the way in which words are put together but in the level of "syntactical organization and subject continuity." It was his minimal contact with any systematic inputs from the home environment and the limited exposure to a variety of stimuli that affected the impoverished child's level of verbal usage.[44]

The notions that an environment deprived of a variety of visual and auditory stimuli affected cognition and that impoverished children were likely to be influenced by the deprivation of such stimuli constituted a popular explanation for the underachievement of millions of American children in the early 1960s. Prominent spokespersons for this position were Martin and Cynthia Deutsch, psychiatrists at New York Medical College, where Martin initiated much of his work on language development.

Fundamental to this argument was the belief that perception was largely dependent upon past experience, taking issue with the Gestalt belief that the internal configuration determined the internal order. The research on animals had indicated that the visual experience had a marked influence on the cognitive behavior of cats, chimpanzees and rats, and studies of auditory and visually deprived humans suggested deterimental effects. Blindfolded college students kept in a soundproof room were observed to have developed an impaired reticular system (which rejected further audivisual stimulation) and lethargic behavior after a day or two.[45] A study of persons blinded at birth by cataracts reported that the subject had an extraordinarily difficult time discriminating even the most simple figures (such as squares and triangles), after the removal of these impairments in adulthood.[46] Correlations were also found between figured sheets, complex crib stabiles and the early psychomotor development of infants.[47]

If it was likely that under certain conditons the presence or absence of certain environmental stimuli could either enhance or curtail an organism's cognitive behavior, what evidence existed by the mid-1960s that the condition of economic poverty deprived children of the necessary sensory input for normal intellectual functioning? Linking underachievement to a culture of poverty's stimulus deprivation was almost entirely speculative, but the widely publicized views of the Deutschs did have modest empirical support. Perhaps Martin's most important research on the subject was a study he concluded with Phyllis Katz (published in 1963) which attempted to assess the relationship between the perceptual and cognitive performances of black lower class boys in Harlem. Good and poor readers at the first, third and fifth grade levels were given a variety of tests to measure perceptivity. The study's major finding was that "normal" readers were superior to "poor" readers in such areas as simple reaction time, shifting auditorv and visual modalities, sustained attention, visual signal detection, and visual and auditory word discrimination. Although there was evidence in the studv that could have been interpreted as suggesting that maturation, rather than environmental condition, may have affected the reported perceptual differences of good and poor readers (the differences generally decreased in the older children), the authors comment on the findings by stating that many poor children mav requrie "specific intervention" to correct culturally determined, poor auditory and visual responses.[48]

According to Deutsch, the urban slum, with its overcrowded living conditions, contained apartments with ". . . few if anv pictures on the wall" and household objects ("be they toys, furniture or utensils," which tended "to be sparse, repetitious and lacking in form and color variations"). He admitted that the paucity of data on the subject mitigated against drawing any conclusions about the effects of "sparsitv of manipulative objects on visual perception." However, Deutsch considered it "an important area, as among the skills necessarv for reading are form discrimination and visual spatial organization."[49]

Zeroing in on the modality of hearing, Cvnthia Deutsch reported a significant correlation between social class and auditorv discrimination skills among black and white first and fifth graders. She contended, however, that it was not the paucitv of sounds that may restrict auditory discrimination in the slum child, but his saturation by sounds, his chaotic bombardment bv "noise" which mav have caused interference and a "turning out" of auditory stimuli ("learned inattention").[50]

Whatever the nature of cognition's interaction with visual and auditory stimuli, by the mid-1960s there was a belief in some educational circles that middle class children did enjoy an advantage over lower class children (particularlv black children) because the impoverished youth lacked many of the sensorv amenities characteristic of the social milieu of their more affluent counterparts. Without these stimulants, schooling for many lower class children was thought to be a difficult and an increasingly unrewarding experience. Surely Martin Deutsch summed up the feelings of many educators in his introduction to his 1963 article on stimulus deprivation,

> "The thesis here is that the lower-class child enters the school situation so poorly prepared to produce what the school demands that initial failure is almost inevitable, and the school experience becomes negatively rather than positively reinforcing."[51]

Deprived of the proper stimuli, restricted by an unappropriate language and dehumanized by a weak self-concept, the academic failure of millions of black children did seem almost inevitable in the early 1960s. In education, something needed to be done to arrest the tide, and pre-school education, with its emphasis upon the building of adequate cognitive skills, was viewed as perhaps the best means to give black children not only initial school success but also long term cognitive benefits.

The Issue of Critical Periods

One of the most important of the many topics germaine to the effects of the environment on cognition in the early 1960s was the so-called "critical periods" hypothesis. Behavioral scientists and educators, such as Benjamin Bloom, J. McVicker Hunt and Arthur Jensen, seemed to believe that there are certain optimal periods in the development of an organism during which exposure to appropriate experiences or stimuli will bring about learning much more easily than in other periods. In Stability and Change in Human Characteristics, Bloom did not use the term "critical period" but argued nonetheless that it is during the early years, when the human organism is undergoing its greatest change, that the environment has greatest influence

> "Variations in the environment have greatest quantitative effect on a characteristic at its most rapid period of change and least effect on a characteristic during the least rapid period of change."[52]

Bloom argued that it is during childhood that at least half of human intelligence develops (perhaps 40 percent by age four; 80 percent by age eight), and that it would be difficult to make up at a later period whatever effects a "deprived" environment may have had during the critical early years.[53]

> ". . . we have assumed that the loss of development in one period cannot be fully recovered in another . . . that is, deprivation in the first four years of life can have far greater consequences than deprivation in the ten years from age 8 through 17."[54]

Bloom contended that in "extreme" environments an average of 2.5 I.Q. points a year may be lost the first four years of life.[55]

Hunt also implied that those in economic want may have permanent scars ("The differnece between the culturally deprived and culturally privileged is, for children, analogous to the difference between cage-reared and pet-reared rats and dogs"),[56] but contended that retardation which occurs during the second and third year of life "may be reversed to a considerable degree" by proper pre-school treatment along the lines of those proposed by Maria Montessori.[57]

Indeed, so accepted was the critical periods hypothesis in this period that Arthur Jensen, now strongly associated with the hereditarian camp, (see Chapter Six) suggested in 1963 that the early experiences of "children of impoverished culture" had a sustaintial influence on their cognitive retardation

> "Our present knowledge of the development of learning abilities indicates that the pre-school years are the most important years of learning in the child's life. A tremendous amount of learning takes place during these years; and this learning is the foundation for all further learnings."[58]

But what was "the present knowledge of learning abilities" Jensen was referring to, i.e., longitudinal data which supported the contention that the environment had greatest quantitative influence during the first few years of life. Jensen, like many other critical period theorists, cited animal studies, and offered the above analysis after reviewing the lasting chemical and behavioral changes produced by a "nursery school" for rats at the University of California.[59] In the case of humans, Benjamin Bloom, probably the most influential and widely quoted spokesperson for the critical periods hypothesis, essentially based his argument on the dubious findings of only three studies.

In his review of the several large scale longitudinal studies of I.Q., he uses Nancy Bayley's Berkeley Growth Study, the most complete and carefully conducted of the group, as providing the clearest evidence of the rapid growth of intelligence during early childhood (I.Q. correlations at age 17 and 2 is +.41, at age 4 is +.71, at age 11 is +.92). While Bayley's work may tell us a good deal about the stability of measured intelligence after age 4, it by no means gives us evidence which either precludes a hereditarian explanation for the phenomenon or shows that measured intelligence is a more or less fixed characteristic after childhood. Once again, it should be emphasized that there is a strong correlation between I.Q. at ages 11 and 17, perhaps because there is typically very little environmental change during this six-year period.

Bloom then zeroes in on the research of E.S. Lee who measured the I.Q. of several groups of black children who spent varying periods of time living in Philadelphia. According to Lee, the I.Q.s of children who were born and raised in that city changed hardly at all between grades one and nine, but the I.Q. change during this period for black children who moved to Philadelphia (usually from the South) by age six was six points. Children arriving by age nine and eleven showed I.Q. gains by grade nine, of four and two points respectively.[60]

Since Lee only measured I.Q. change up to grade nine, was it appropriate to argue that the environments most profound effect occurs during the early years? The fact that childrens' I.Q.s changed six points in nine years (grade one through nine), four points in five years (age 9 or fourth grade through ninth grade) and two points in three years (age 11 or sixth grade through ninth grade) may only tell us that the longer the environmental exposure the greater the change. If those children who arrived at ages nine and eleven had been followed for nine years (to ages 18 and 21), Lee might have found greater than 4 and 2 point changes in their I.Q. If he had followed all children for at least nine years and found the greatest changes in the youngest children, one could then, and only then, conclude that the data suggested the significance of early environmental exposure.

The third report Bloom cites he considered "most crucial in establishing the pattern of change in relation to the environment." In this study James Kirk followed eighty-one mentally retarded children (I.Q.s 45-80) three to six year olds over a period covering three to five years. Both the experimental and contrast groups were composed of children who lived in a regular community and children who resided at institutions for the retarded. The experimental subjects attended a pre-school either in the community or at an institution; the control subjects did not. Kirk found that the I.Q.s of most experimental subjects increased during their pre-school experience (while the I.Q.s of the controls did not) and that these gains did not decline but were generally maintained after the treatment. But what Bloom does not include in his interpretation is that the I.Q.s of most children in the community contrast group, upon entering first grade or a special class at age six, increased to a level "approaching the score of the experimental children who had attended pre-school." Commenting on this unexpected finding, Kirk questions the significance of pre-schools for retarded children

> "If these results are corroborated by later studies, it could mean that pre-schools for mentally handicapped children are not necessary, since the children will accelerate their rate of development after entering school at the unusual age of six."[61]

Kirk does argue, however, for pre-school for children from "psychosocially deprived" homes, after analyzing his finding that the subjects' siblings from such environments did not show the characteristic I.Q. gain in the first grade. But his research also suggests that under certain conditions I.Q. can be increased as easily in the first grade as it can in pre-school, which not only challenges the "earlier the better" assumptions but questions the irreversibility of early childhood experience.[62]

If the three studies emphasized by Bloom offer less than convincing support for the critical periods hypothesis, was there relevant research which was not cited in Stability and Change in Human Characteristics? A 1952 review of the literature on maternal deprivation conducted by John Bowlby for the United Nations suggested that a severely deprived early environment may do irreparable damage affectively. The study offered very little longitudinal data indicating any sustained effect on intellect.[63] A 1962 study by Sally Province, reporting the physical, emotion and cognitive retardation of institutionalized infants, also found them generally below the norms on the various measures when tested shortly after placement in foster homes. But Province commented that "we have been impressed by and filled with admiration for the adaptability, resiliency, and capacity for improvement we have witnessed in the children in the course of this research."[64]

An interesting 1960 analysis by Wayne Dennis of the long term effects of infant crib confinement of Iranian orphans has remained virtually ignored. These orphans spent up to 24 hours a day for "many months" during their first year in the supine position. Needless to say, the children enjoyed very little stimulation (they were even fed by a bottle resting on a pillow). In this situation, and at age three only 15 percent could walk alone. At approximately three years of age, however, many of these children were moved to another institution which offered them substantially greater environmental enrichment. In the second institution Dennis observed the behavior of children between the ages of 6 and 16 who were only similarly confined and ". . . presumably as retarded at ages two or three" as were the children whose behavior was described above. However, Dennis found these children

> ". . .attending school, playing games, doing chores and being trained in different skills, such as the weaving of Persian rugs. There was nothing in their general behavior to suggest that any permanent consequences ensued from the extreme retardation in motor development during the early years."[65]

The study which may have offered the strongest evidence to support the significance of early childhood was conducted by William Goldfarb. Published in 1943, it was for some reason given very little attention by critical period theorists of the early 1960s (Hunt gives it only the briefest reference near the end of Intelligence and Experience; Bloom does not cite it at all in Stability and Change in Human Characteristics). In any event, Goldfarb reported highly significant differences at age 12 years in I.Q. and achievement between a group of fifteen foster children who had spent most of their first two years of life in institutions and another group of foster children, equated by number, age, sex and foster maternal background, who had always lived with "real families." The fact that Goldfarb reported on such a small sample and his omission of any explanation of the subject selection process may have mitigated against his study enjoying greater publicity.[66]

In general, the critical periods theory of the early 1960s was more a statement of faith than a working hypothesis. The longitudinal studies such as Bayley's only gave us age level correlational data on intelligence; the Lee, Kirk and Dennis reports suggested that I.Q.s may increase significantly at age six and that psychomotor normality can emerge following infant confinement. The work of Kirk and Goldfarb does offer some evidence for the sustenance of early learning experience, but the extensive Bowlby review suggests that it may have a greater effect on affective change, rather than cognitive growth.

In short, by 1965 much of the research on cognition suggested that the environment could have a profound effect on measured intelligence and school achievement, but it was unclear what cultural variables may have contributed to academic retardation and whether there was any optimal time in human development for receptivity to learning. Heredity seemed to account for most of the variation in I.Q. among a homogeneous population, but the environment took on a greater dimension in explaining the I.Q. differences between people reared in contrasting environments. In the early 1960s it was fashionable to describe the contrasting environment of the poor as deprived of proper self-concept, appropriate language and sensory stimuli, and that variables such as these within a deprived culture could prevent the natural unfolding of one's genetic potential. Did, however, the sub-culture of poverty of many black Americans actually depress cognition or did it simply develop different cognitive processes which were inadequately measured on culturally biased intelligence tests?

The question of whether the black sub-culture is a deficient or only a different environment will be discussed in Chapter Nine. For the moment, it is sufficient to conclude that the theories and correlational research findings suggested that the socio-cultural milieu of black children mitigated against normal intellectual growth, at least as defined by the dominant American culture. Accordingly, if these inadequate environmental conditions that inhibited measured intelligence and school achievement could be somehow compensated for by exposing disadvantaged blacks to the cultural variables strongly associated with intelligence. it should be possible to significantly reduce the inequality in cognitive achievement which existed between the white and black American population. By the early 1960s, a significant segment of the American population had come to regard the schools as the institution that could successfully provide that compensation.

The Schools and Cognition

If the attitude that the environment has a profound influence on cognition has historically enjoyed varying degrees of acceptance, there is little question that the assumption that the schools constitute a very important part of the learning has long been apart of the American ethos. Horace Mann's famous labeling of education as "the great equalizer" in 1847,[67] the unprecedented growth of public education in this country during the 19th century, and the compulsory attendance laws in existance in most states by 1900 are indicative of the value placed on the school's ability to offer millions of people, often from diverse cultural backgrounds, greater equality of opportunity.

The relevant research also supports the significance of formal education. Many elementary schools in Holland were closed for a prolonged period during World War II, and a reported I.Q. decline of some Dutch children at that time has been attributed to the absence of public education.[68] In Prince Edward County and in New York City the schools were closed for varying periods during the sixties because of respective battles over desegregation and teacher contracts. In both situations children's achievement was apparently significantly affected by the closings.[69]

With the advent of the Civil Rights Movement the schools, primarily through desegregation and compensatory education, were seen as playing a crucial role in the struggle to further democratize American Society politically, socially and economically. The rhetoric of the period on compensatory education may be summarized by drawing on the statements of influential behavioral scientists and educators attending two conferences on the problems of disadvantaged students: one sponsored by Teachers College, Columbia University in 1962 and another by the University of Chicago in 1964.

Under a grant from the Ford Foundation, the Columbia conference invited scholars from twenty-four cities to New York in July of 1962 ". . . to examine the many dimensions of education in depressed urban areas and to develop sound guiding principles for program planners in city school systems."[70] According to Columbia's A. Harry Passow, the conference coordinator, "over and over, participants stuck on the question of whether the school can make the necessary impact without society really equalizing opportunities in employment, in housing, in civic affairs."[71] But it was unmistakable that the school was viewed as capable of effectively compensating for much of the depressing effects of urban poverty. Mel Ravitz, an Associate Professor of sociology and anthropology from Wayne State University concluded:

> "Major stress is placed on the school to broaden the horizons of children because often parents do not care or are unable to do very much to enrich their children's experience. . .

> The key figure in the entire educational process is the teacher. Good teachers can work miracles with children coming from any background. . ."[72]

Referring to the perpetuating effects of poverty on "the general academic inadequacy of the majority of disadvantaged pupils. . ." Miriam Golberg of Teachers College suggested:

> "At some point the circular negative reinforcement has to be attacked. Perhaps the most accessible place is the school itself. One of the major issues confronting education today is to discover the means by which the school can compensate for the lack of readiness for learning which lower class children, in general, and the Negro and other discriminated-against groups of children, in particular, bring to their school work."[73]

Martin Deutsch saw the school as being able to ". . . significantly reduce the attenuating influence of the socially marginal environment" and Kenneth Clark professed considerable faith in education as a viable agent of social change.

> "Education has been one of the most effective means for social mobility in the American society. This problem in the future may be different from the similar problems in the past only in that it will involve different and larger groups of previously disadvantaged individuals."[74]

The second conference (the "Research Conference on Education and Cultural Deprivation") was held at the University of Chicago in June, 1964 to review ". . . the problems of educational and cultural deprivation, to make recommendations to solve some of the problems, and to suggest problems for further research. Funded by the U.S. Office of Education, the conference solicited working papers from thirty-one scholars from a variety of disciplines. Among the conference participants were Basil Bernstein, Benjamin Bloom, Martin Deutsch, Erik Erickson, Arthur Jensen, Lawrence Kohlberg and Thomas Pettigrew.[75] A summary of the conference findings, by Benjamin Bloom, Allison Davis and Robert Hess, should serve to capture the flavor of the attitudes expressed in the many meetings and papers. After reflecting upon the wisdom and extraordinary difficulty of revolutionizing American education as a whole to adjust to an increasingly complex and rapidly changing world, the authors address the specific problems of the "culturally deprived."

> "But, there is a much more immediate problem. This is in some ways an easier problem to attack and it must be solved in the present. We cannot wait for a decade in which to gradually find solutions for this problem.
>
> In the present educational system in the U.S. (and elsewhere) we find a substantial group of students who do not make normal progress in their school learning.
>
> . . . .We will refer to this group as culturally disadvantaged or culturally deprived because we believe the roots of their problem may in large part be traced to their experiences in homes which do not transmit the cultural patterns necessary for the types of learning characteristic of the schools and the larger society."[76]

Following a description of the nature of cultural deprivation, the conference turned to the educational alternative.

> "What is needed to solve our current as well as future crisis in education is a system of compensatory education which can prevent or overcome earlier deficiencies in the development of each individual. Essentially, what this involves is the writing and filling of educational prescriptions for groups of children which will enable them to realize their fullest development."[77]

If there was a prevailing feeling among educators that a system of compensatory education could prevent or overcome environmentally determined deficiencies, what specific educational prescriptions should be offered that traditional treatment had failed to cure. There was, of course, no consensus. There was a paucity of research on the subject and few successful models to follow. These factors combined with the general idealism of the early 1960s contributed to what was probably a period of unprecendented experimentation in American education. But many proponents of compensatory education felt they were not entirely shooting in darkness. There was some psychological and sociological understanding of the relationship between poverty and emergent from the interpretations of this literature was a myriad of suggestions for educational reform that were divergent, yet typically contained certain common denominators.

At the pre-school level a great emphasis was placed on the use of language. Whether one is reading the teacher-directed behavioral technique of Carl Bereiter and Sigfreid Engelmann, the cognitive-discovery approaches of Merle Karnes, David Weikart, and Martin Deutsch or the more traditional pre-school curriculum of Carl Beller or the Howard University Project, the importance of the child's verbal interaction with adults is stressed. The suggestions of Martin Deutsch in 1964 were typical of those advocating more academic, structured pre-school programs:

> A language training program would require the creation of a rich, individualized language environment, where words are repeatedly placed in a meaningful context, and where the child is allowed multiple opportunities for expressive language demonstrations as well as for receiving language stimuli under optimal conditions and being encouraged to make appropriate responses.[78]

According to Deutsch, an enrichment language program would improve significantly the pre-school child's preparation for academic success in the early elementary grades thereby reducing the chances of motivational problems stemming from frequent failure.[79]

At the elementary and secondary school level it was often suggested that the schools work directly to improve the child's self-image--"building in these children," according to Kenneth Clark, "a positive self-esteem to supplant the feeling of inferiority and sense of hopelessness which are supported by the all too-pervasive pattern of social realities."[80] Strategies for improving self-concepts included the use of textbooks and materials which reflected to a greater degree the life experiences of disadvantaged minority children, the employment of teachers familiar with the culture and ethnic background of their pupils and a generous use of positive reinforcement, often in terms of verbal praise.[81]

Reforms in the administration and curriculum of the public elementary and secondary schools were suggested in the "Great Cities School Improvement Studies," of the late 1950s, sponsored by the Ford Foundation.

> 1. An extended school day and school week (to include field trips to civic, recreational, industrial and other centers of interest, as well as reading clinics, opportunities for recreational reading in the school library, small academic coaching and small group guidance) to improve the basic skills, motivation and prevent attrition.
>
> 2. A flexible, non-graded grouping of children in the elementary school.

> 3. The organization of centers or classes to provide a special program for the culturally deprived child who is of high school age. . .
>
> 4. Varying the sizes of classes within the school day (so that the particular talents of some teachers are brought to large groups of children, and at the same time other teachers, who may have talents for working with the culturally deprived in small groups or as individuals, are freed for work of this type. . .
>
> 5. An organization in which the length of the periods in the school day is altered to give the culturally deprived child some short periods of instruction in small groups in skill areas and longer periods of integrated unit activity will improve reading and arithmetic skills, establish the close, stabilizing relationship with an adult. . .
>
> 6. Greater use of paraprofessionals. . .for such purposes as vision and hearing screening, . . . field trips, operating projectors. . . will release professional personnel for teaching purposes and improve the academic achievement of the culturally deprived . . .[82]

Educational reforms called for by Patricia Sexton in her 1961 publication Education and Income included an elimination of "segregated groupings and curriculum," a "replacement of the highly competitive system of marks, exams and comparisons of all sorts . . . by other types of incentives to learning, the removal of irrelevant 'dead weight' from the curriculum such as 'meaningless dates and data,' an expansion of work-study programs, more attractive accessible libraries, and greater 'attention' to an unexplainably neglected skill concentration. . ."[83]

Effective teachers for disadvantaged children were viewed by Sexton as perhaps more likely to be male ("efforts should be made to encourage more men, of the type boys can readily identify with, to enter teaching"), enthusiastic, and well paid; by Miriam Goldberg as respectful, familiar with the child's cultural experience, sensitive to the "self-fulfilling prophecy," capable of showmanship, and by David Riessman In the Culturally Deprived Child (1962) as consistent, straight-forward, down-to-earth, sometimes physical, and dedicated.[84]

The suggested changes just cited in administration, curriculum, and teaching represent only a microscopic glance at the proposed reforms in the schools to improve the cognitive achievement of disadvantaged pupils. Nevertheless, they do give us a "feel" for the ideal characteristics of many, if not most, of the proposed compensatory education programs of the 1960s. Drawing on the suggested reforms listed in this section and in many additional discussions of the subject, the following listing includes some of the more commonly recommended components of early compensatory education programs:

1. Improving the motivation of pupils.
2. Improving the self-concepts of pupils.
3. Increasing per pupil expenditures.
4. Lowering the teacher-pupil ratio.
5. Individualizing instruction where appropriate.
6. Liberalizing the administration of the schools to facilitate appropriate changes in class periods, class size, and student and teacher mobility.
7. Ordering or creating instructional materials which are more consistent with the socio-cultural background of the pupils.
8. Emphasizing "language training" at the pre-school level.
9. Employing enthusiastic and dedicated teachers who are familiar with the pupils' socio-cultural background, understand the nature of a "disadvantaged" environment's effect on cognitive achievement, and think positively about the school's power to effectively compensate for environmental inadequacies.
10. Involving members of the community--particularly the pupils parents--in the children's learning process.

In the early 1960s compensatory education may have been defined, therefore, as a process whereby educators attempted to compensate for the academic inadequacies of economically and/or socially disadvantaged children by giving particular attention to the strengthening of basic cognitive skills, motivation and self-concept in an atmosphere imbued with the promotion of individual needs by employing positive reinforcement, cultural enrichment, student-centered materials and humanistic teaching.

EARLY ELEMENTARY AND SECONDARY PROGRAMS

The first compensatory education program to gain national prominence began in a Harlem junior high in 1956.[85] Entitled the Demonstration Guidance Project, it sought to identify pupils with promising standardized test credentials (above the 50th percentile) and provide them with intensive counseling and individualized teaching.[86] After impressive gains were reported in achievement,* the program was expanded in 1959 to a large-scale project entitled Higher Horizons which provided comprehensive compensatory services to thousands of disadvantaged low achievers in the Harlem area. In 1967 the U.S. Commission on Civil Rights described the project as the largest program of its kind in American history ". . . involving by 1962, 64,000 children from 52 elementary schools, 13 junior high schools and 2 senior high schools."[87]

Higher Horizons was not only the largest compensatory education program of the early 1960s but probably the most influential. In 1963 Passow described the project as ". . . perhaps the most widely known enrichment program. . . now being adopted in numerous other communities" and in 1969 White House aide Roger Freemen contended that "during the early 1960s Higher Horizons was widely promised as a shining example and was copied in many cities."[88] Indeed, it was used as a model for DHEW's shaping of proposals for Title I ESEA.[89]

According to the U.S. Commission on Civil Rights

> "Four major techniques were used on Higher Horizons. First, teachers were trained and encouraged to improve both their expectations of the students and their own ability to teach disadvantaged children. Second, counseling and guidance services were extended and increased to raise aspirations. . . Third, an effort was made to broaden cultural backgrounds and horizons of students through visits to museums, libraries, colleges and concerts. Lastly Special remedial teachers were provided to upgrade reading, writing and arithmetic skills.[90]

Another well-publicized general enrichment program was Project Bannecker in St. Louis, initiated in 1957 under the energetic, vocal leadership of Samuel Shepard. By 1965 the project had become one of the nation's largest compensatory programs involving twenty-three predominately black elementary schools with more than 14,000 pupils.[92] The Bannecker project made few curriculum changes and placed great emphasis on such affective and community modifications as improving the motivation and self-concepts of pupils, bettering the attitudes of teachers (often by "re-education,") and increasing the involvement of parents. Interschool academic competition was staged, pep rallies were held and staff home visits were encouraged.[93]

---

*For 147 of 250 pupils, a 4.3 years reading gain after 2.6 years of the program.[91]

In addition to the influential pioneering efforts in New York and St. Louis, scores of enrichment programs emerged in many other American cities and towns.* Among those receiving the greatest attention were sponsored by the Ford Foundations Great City Improvement Project, a comprehensive educational enrichment project beginning in 1957 targeted at the urban poor in Milwaukee, Berkeley, Pittsburg, St. Louis, Washington, D.C., Cleveland and Detroit. Although the programs varied considerably from city to city, Dorsey Baynham, a free lance education writer, wrote in 1963 that three factors common to each site were

1. A "willingness to experiment with a broad range of teaching materials. . .and. . .administrative approaches"
2. "Strenuous efforts to search out and use community help..." and
3. Preparation of teachers to meet the special affective and cognitive needs of disadvantaged children.[94]

Early Pre-School Programs and Head Start

At the pre-school level, the major impetus for compensatory education both theorectically and practically came not from the school systems or social agencies but from a few academicians from the universities and colleges.[95] In the early 1960s compensatory pre-school programs may not have exceeded two dozen,[96] and most of them were experimental research projects. Among those that received the greatest publicity and employed the most sophisticated research designs were Martin Deutsch's New York based Initial Enrichment Program, (one of the earliest, beginning in 1958),

---

*An exact number of pre-ESEA compensatory programs is impossible to determine. A 1965 inventory by the University of Chicago's Urban Child Center listed several hundred programs.

Merle Karnes' Ameliorative Pre-School Program at the University of Illinois (a cognitive oriented whole child model), Susan Gray and Rupert Klaus's Early Training Project at George Peabody College in Nashville (another cognitive "whole child" program), David Weikart's Piagetian Ypsilanti Perry Pre-School Project in Michigan and Bereiter-Engelmann's behavioristic design at the University of Illinois.[97]

The public sector made a rather dramatic entry into the pre-school business, however, when it began serving through Project Head Start 561,000 children in 2,400 centers in the summer of 1965. Project Head Start was initiated by the War on Poverty's Economic Opportunity Act of 1964. Community action agencies, school systems and various voluntary organizations were eligible for sponsorship.[98]

Early childhood educator Millie Almy has described Head Start as a project that from its inception "captured the imagination of many people" and had originally the "rather modest and realistic goal" of developing "pleasure, confidence and courage" in disadvantaged pre-schoolers. She observes that "The scope of the program grew, however, to include an educational program, health services..., social services, psychological services, nutrition and a parent participation program." Despite its modest beginning, the focus of assessments of Head Start effectiveness, as they were for the university pre-school models, soon became cognitive growth (see Chapter Five).

Title I

Even more encompassing than Project Head Start is the Elementary and Secondary Education Act (ESEA), which was passed in 1965 after a lengthy congressional debate over the wisdom of federal aid to education. ESEA represented the first legislation which contained a large-scale commitment to use federal funds to improve the equality of American education,[99] and Title I of the act was specifically written to finance compensatory programs for an academically disadvantaged school age population. One may appreciate the relative size of the monies appropriated by recognizing that the 1.2 billion dollars provided by Congress under Title I the first year exceeded "the total Office of Education direct appropriations" for the entire year of 1964.[100]

Although ESEA's many Titles monetarily supported such educational components as libraries, innovative education projects and civil rights compliance, its success in compensating for environmentally caused academic retardation soon became a primary measure of ESEA effectiveness. For, in the words of one close observer, in addition to its being the first large federal education act, ESEA's ". . . second unique feature was . . . the evaluation requirement included in Title I," which marked ". . . the first time that federal legislation included an evaluation requirement within the law itself."[101] This requirement became the impetus for annual evaluations of Title I at the national, state and local levels and eventually nationwide searches for exemplary programs for proof that ESEA Title I was "working." What generally came to be meant by working was "hard" data showing that programs were producing average pupil reading or math achievement gains of at least a month of instruction (1:1), i.e., at the same growth rate or better than the national average.[102]

CRITIQUING THE RATIONALE

Despite the cautiously realistic goals of many Head Start founders and the reservations expressed about the powers of schooling by many academicians, there appeared to be widespread acceptance of the notion that improved schooling could indeed compensate for the environmentally determined deficiencies in measured intelligence and pupil achievement.

What factors could have contributed to such an idealistic criteria by which to assess the success of effectiveness of pre-school, elementary schools and secondary school compensatory programs? There were only a handful of studies showing significant pupil gains in compensatory situations,* a promising first year evaluation of Higher Horizons, a Bloom analysis of the Chicago Group twin data suggesting that "educational" advantage had a greater influence on I.Q. than "environmental" advantages.[103] But a careful reading of Dr. Bloom's argument on education and twins, just as in his argument for critical periods, will raise questions about the validity of his conclusions.

---

*

If the research was less than convincing, the common rationale for compensatory education was also of questionable logic. If we list the assumptions summarizing the rationale offered at the beginning of this chapter as premises, the illogic becomes apparent.

MAJOR PREMISE:	The environment has a profound influence on measured intelligence and school achievement.
MINOR PREMISE:	The schools constitute an important part of the environment.
CONCLUSION:	Schools can compensate for the environmentally determined defiencies in measured intelligence and school achievement.

Assuming the validity of the major and minor premises, it is theorectically impossible for any part of the whole to equal the whole. The word "compensate," as defined by Merriam-Webster, means "to be equivalent to" or "to supply an equivalent."[104] Consequently, it would be logical to argue that "the schools can have an important influence on measured intelligence and school achievement," but illogical to argue they can wholly compensate for environmentally determined deficiencies."

Logic aside, in their recognition of the significance of the environment's influence, compensatory educators of the early sixties recommended schooling which would expose children to much more than the so-called "basics." Cultural enrichment activities such as trips to museums, zoos and concerts were often recommended for disadvantaged children. It is questionable, however, whether formalizing these experiences a few hours a week through the schools could offer an adequate substitute for the totality of the child's socio-cultural milieu. If educationally significant gains in reading and mathematics were the primary goals of the early programs, it may have been better to have devoted greater attention to the specific teaching of these skills.

The fact that the so-called "culturally deprived" model was adopted by many early compensatory programs may have compounded the problem. Exposing children to the cognitive processess of a given culture to facilitate the comprehension of its dialect may be educationally sound practice, no matter how limited the immersion. Designing for children ego inflating experiences, language training activities and multi-sensory stimulations, however, may be another matter. Not only was such an approach potentially dehumanizing, but it had very little scientific justification. In the case of blacks, there was no solid evidence that they had lower general self-concepts than whites, only speculation that they spoke a restricted tongue and no empirical data suggesting that their physical environment was deprived of intellect-building stimuli. The fact that the black sub-culture was different did not necessarily mean it was deprived. It may have ill prepared its members to score well on middle class measures of intelligence, but may not have inhibited cognition in any absolute manner. Calling black children "culturally deprived" probably did little to enhance their self-esteem, and advocating for urban pre-schoolers language training (as if they had been reared in attics or in cribs) as a pre-requisite for normal achievement in reading may have unappropriately taxed the time and energy of many early compensatory education programs.

COLEMAN ADDS TO THE CONTROVERSY

Whatever may have been the inherent quality of the black way of life in 1965, there is no question that the influences of slavery, segregation, prejudice and poverty had ill-prepared millions of black children for equality of achievement opportunities. For many other minority and white children prejudices and poverty had also taken its toll. Because of a growing concern about education's relation to income and ethnicity, the year 1965 not only encompassed the origin of Head Start, the passage of ESEA but the initiation of the massive Equality of Education Opportunity Survey (the Coleman Report), directed by James Coleman.[105] The Coleman Report was initiated in response to Section 402 of the Civil Rights Act of 1964 which ordered that the Commissioner of Education ". . . conduct a survey and make a report to the President and the Congress. . . concerning the lack of availability of equal opportunities for individuals by reason of race, color, religion or national origin in public educational institutions at all levels in the United States. . ."[106]

Called by Mosteller and Moynihan ". . . the second largest social science research project in history "(Project Talent it seems was larger),[107] the survey tested some 570,000 pupils and 60,000 teachers.[108] Data from over 4,000 schools was collected and analyzed in extraordinary detail.[109] To Coleman and his staff equality of educational opportunity apparently meant not only equalizing inputs (school facilities and per pupil expenditure) but school outputs (pupil achievement on standardized tests), for their report went well beyond attempts to document the limited school resources generally available to racial and ethnic minorities.[110]

The Coleman Report is notable not for what it found but for what it did not find. To the surprise of almost everyone the survey did not report gross inequalities in educational resources in schools with differing minority enrollment and did not find evidence that school facilities and curriculum in themselves had much at all to do with pupil achievement.[111] Nationally, whites did enjoy a greater quantity of school resources which were thought to affect learning, (than did blacks) but when available facilities for the two races were analyzed regionally (focusing on the South,[112] Midwest, etc.), remarkably little difference in educational services was reported.[113] The input variables of "teacher characteristics" (p. 316) and "student body characteristics" did correlate with measured pupil learning, but

> Differences in school facilities and curriculum, which are the major variables by which attempts are made to improve schools, are so little related to differences in achievement levels of students that, with very few exceptions, their effects fail to appear in a survey of this magnitude.[114]

On the standardized tests used by Coleman, minority children (with the exception of "Oriental Americans") scored well below "majority" children at each grade level reported in the survey,* and there existed a cumulative deficit in achievement between each of the minorities and the majority.[115] The achievement test scores of black children were the lowest of any racial or ethnic group measured, falling about one standard deviation below the white average. Certain teacher characteristics (particularly teacher verbal facility and educational background) were shown as likely to have some effect on black (but not white) achievement,[116] and the student's expressed feeling of destiny control had a strong relationship to achievement** ("stronger. . . than do all the 'school' factors together,").[117] Generally speaking, the percentage of white pupils in a school improved the achievement of Negro pupils but only consistently when white were more than half the school population.[118] Finally, as one might expect, the school factors affecting achievement usually had a greater influence on minority students than on majority pupils.[119]

But even factors such as racial integration and the teacher's verbal facility showed a relatively small correlation with achievement when compared to student environmental background which usually accounted for between 10-25 percent of the variance in individual achievement.[120] Eight background factors were included in the survey dealing with family structure and size, parental economic and educational backgrounds, parental attitudes toward education, availability of reading materials and certain technical amenities, and length of residence in an urban area. Collectively, the pupil background characteristics had a slightly greater effect on the variance in achievement at the earlier grades than the later grades* and a somewhat greater effect on white achievement than on black achievement.[121] The significance of family background compared to the other variables and the importance of the Coleman Report itself may be summarized by quoting the concluding paragraph from the 106 page section on "Pupil Achievement and Motivation."

---

* Grades 1, 3, 6, 9 and 12.

** The attitude of "sense of control of the environment" was highly related to achievement, but this feeling was not found to be influenced much by any school characteristics.

> Taking all these results together, one implication stands out above all: that schools bring little influence to bear on a child's achievement that is independent of his background and general social context; and that this very lack of an independent effect means that the inequalities imposed on children by their home, neighborhood, and peer environment are carried along to become the inequalities with which they confront adult life at the end of school. For equality of educational opportunity through the schools must imply a strong effect of schools that is independent of the child's immediate social environment, and that strong independent effect is not present in American schools.[122]

The "independent effect of school" that was so critical and was found by Coleman to be so inconspicuous in the America of 1965 would be made more visible, it was hoped, by federal appropriation throughout the nation for compensatory education programs. Indeed, by the time the Office of Education had released Equality of Education Opportunity Survey over one billion dollars had already been spent under Title I of ESEA, essentially to permit the schools to effectively educate children independent of their home environment.[123]

The Coleman Report has had far reaching implications for desegregation policies, compensatory education and educational research. (See Chapter Five). For the moment it is sufficient to say that various reanalyses have both generally supported [124] and strongly taken issue with [125] his conclusions, and that more recent smaller scale studies have found that variation in schooling does affect variation in learning.

---

*Only grades 6, 9 and 12 were included. Information on pupil background was obtained from the children themselves; it was believed difficult to question younger children on this topic. But it is interesting that variation in schooling had a greater effect in the higher grades than in the lower grades.

Perhaps Coleman's greatest contribution to educational sociology is not, therefore, his conclusion that schools make little difference, but his generation of thought and research on just how much difference schooling can make. The evaluations reviewed in Chapter Five and the nature-nurture discussion in Chapter Six will hopefully give us a better insight into the answer to this question. We may find the data suggesting that schools can improve the achievement of disadvantaged children but not compensate for environmental determined deficiencies. If this is the case, greater attention could be given to the inherent illogic of compensatory education's rationale. On the other hand, we may find data suggesting the schools are capable of such compensation and that will then raise questions about the entire environmental hypothesis. Or perhaps in this society it is education, not socio-cultural background, which tells us most about pupil achievement, and the impact of schooling is potentially so powerful that appropriate instruction can overcome what little influence the out-of-school environment may have. Another possibility is that genetic endowment is the primary cause of the gross inequalities in educational achievement and if this is the case, we must ask what the schools can do, if anything, to neutralize the effects of heredity.

Part II

THE EFFECTIVENESS OF COMPENSATORY EDUCATION: WHAT THE RESEARCH SHOWS

If the period from 1945-1965 can be described as providing a rationale for integration and compensatory education policies and as contributing to the development of a consensus on what was appropriate educationally for black Americans, the years since 1965 may be logically referred to as a period of evalution of these strategies and as cultivating considerable dissension among many reformist educators and academicians. By 1955 school desegregation had begun to be ordered, by 1965 Title I had been legislated and by that mid-sixties summer a President from Texas was overseeing the enrollment of approximately 500,000children in Head Start. But the release of the Coleman Report in 1966 and the onslaught of educational assessments which followed soon ignited continuous controversies surrounding the explosive issues of race and education. Indeed by 1968 Head Start was judged by some observers to be ineffective, by 1969 a former proponent of compensatory education was offering a possible genetic explanation for the "apparent failure" of compensatory education and by 1975 Dr. Coleman, a man who formerly enjoyed impeccable liberal credentials, was under attack for yet another report on the alleged ineffectiveness of desegregation due to "white flight."

Coleman's first report, the massive Equality of Educational Opportunity Survey (EEOS), was perhaps not as inflamatory as his conclusions on busing, but was surely every bit as provocative. His finding that black achievement was related to the student body's racial composition generally furthered the interests of integration in general and busing in particular.[1] However, the fact that white pupil presence correlated only very modestly with black gains often contributed to a questioning of the necessity of racial balance in the schools. Coleman's data was gathered too early* to measure any effects of Title I, but his conclusions regarding the relative ineffectiveness of schooling in equalizing achievement triggered a number of reanalyses of his survey. In the fall of 1966, a Harvard hosted seminar entitled "On Equality of Educational Opportunity," was described by Daniel Patrick Moynihan as taking on "near conference proportions" and as instrumental in warranting the creation of an entire educational policy center at Harvard to begin a more sophisticated EEOS re-examination.[2]

---

*The Fall of 1965

In the field of compensatory education, a U.S. Commission on Civil Rights study of pre-Title I programs appeared in 1967 and a longitudinal study of Head Start was released in 1969. By far the greatest contributor to evaluation quantity, however, was that unique Title I feature which mandated objective assessments of the programs impact. In the five year period following ESEA passage, over one hundred of the myriad of Title I state and local evaluations found their way into the public records. In the late 1960s extensive searches for exemplary compensatory education programs were conducted and their findings released and research to identify and publicize compensatory models continued generally unabated throughout the 1970s.

Collectively, the pre-school and public school evalutions have reported mixed findings which have only served to foster moderate political and intense academic debate over how effective schools can be and how best they can be effective. During this period, early childhood programs have been judged effective, then ineffective and finally successful; compensatory education in general has been seen as having failed and then having "worked;" and "back to the basics" proponents and "whole child" advocates have nearly come to blows on such topics as research design, the validity of standardized testing and the appropriate curriculum for increasing and sustaining pupil growth.

This chapter will summarize the relevant research by reviewing the major studies of compensatory education conducted at the national and local levels from 1965 to 1975 and then from 1975 on when the findings became somewhat more encouraging. As contradictory and controversial as the findings have proven to be, it is proposed here that a review of the more scientifically promising evaluations of compensatory education can give us an insight into potential of schooling to redress the racial differences in achievement in America and what may be the most productive means of increasing the effectiveness of the public schools for underachieving black pupils.

NATIONAL EVALUATIONS TO 1975

Credit has been given to Senator Robert Kennedy, more than any other person, for the evaluation requirement being a part of Title I of the Elementary and Secondary Education Act (ESEA). Apparently Kennedy's interest in the evaluation component was based largely on his concern that poor people should be kept abreast of whether tax dollars were being spent wisely.[3] Accordingly, Kennedy proposed some sort of testing or reporting system as an accountability measure. Although Kennedy's enthusiasm for the bill was contingent upon it including an evalution component (an enthusiasm that was thought critical to the Act's passage), the language of the evaluation section was much more general than the New York Senator had wanted.[4] According to the Rand Corporation's Milbrey McLaughlin

> "Kennedy's support was important to the passage of ESEA, but evaluation was also a traditional bugaboo of schoolmen. Thus to appease Senator Kennedy and not simultaneously anger educational interest groups, drafters resorted to additional political diplomacy. Kennedy's demand for an accountability measure was met as inconspicuously as possible, with a lonely worded valuation mandate."[5]

One top Congressional liaison at the United States Office of Education (USOE) explained that when Kennedy's evaluation plan was put into Congressional language, "the guiding concern was that the [evaluation] amendment be broad and general, and open to multiple interpretations at the local level." Consequently, the localities were told only that ". . . provisions for appropriate objective measures of education, will be adopted for evaluation, at least annually on the effectiveness of the programs in meeting the special educational needs of educationally deprived children."[6]

Because of the vague evaluation requirement of the Act and the need to maintain considerable local autonomy, a loosely structured "three-tiered" Title I reporting system emerged.[7] Typically, the Local Educational Agency (LEA), which constituted the third tier, would simply report its "objective" test results to the State Educational Agency (SEA) and a compilation of the various LEA data by this second tier would be passed on to the Commissioner of Education. Given this lowest to highest data collection system, the evaluation inexperience of the USOE staff and the "severe time restraints" facing first tier personnel in their writing and implementation of evaluation guidelines,[8] it should come as no surprise that the earliest national reports on Title I effectiveness produced data of little scientific value. The first year report was published in 1966 and simply summarized state descriptions of their programs, most of which had only been in operation 3 or 4 months.[9] A second annual report appeared in 1967 and once again contained little more than a review of information submitted by the state. (The state data found ". . . many Title I youngsters . . . improving, sometimes gaining a full month for every month spent in the classroom." The document did mention, however, that reports from the cities were "disheartening" with little evidence of significant pupil gains.[10]

It is McLaughlin's observation that the cumbersome three tier data collection method and the suspected "faculty manipulation of the limited data available" soon came under Congressional attack for their contributing to empirical imprecision. According to Kathryn Hecht, one of the USOE pioneers in the area of Title I evaluation, it was primarily Congress, not USOE personnel, that pressured the infant Title I evaluation staff to begin employing achievement test scores as a principal criteria to assess Title I effectiveness.

> "When it became obvious that aggragating locally defined evaluations was not going to produce data to convince Congress Title I was working, the emphasis on achievement tests and aggrating more uniform test scores became an evalution emphasis."[11]

Hecht reminds us that the logic for original data gathering process was "quite consistent with the language of the Act, in that local districts were to determine Title I needs and objectives and were to plan their evalutions relevant to those objectives." In keeping with empirical methodology but hardly with the democratization process, "Congress mandated using reading and math scores as effectiveness criteria regardless of their relationship to the program content."[12]

Accordingly, in 1967 Congress ordered that more sophisticated reports be made by the Commissioner each year on the achievement of Title I pupils. The 1968 survey sampled 465 of 10,544 districts receiving Title I funds. In order to facilitate the collection and analysis of the data, only grades 2, 4 and 6 were included in the study. USOE's interest in obtaining accurate information was indicated by its elimination of all reading achievement data that did not include such components as standardized pre- and post-tests. Consequently, only the reading scores of 11,490 pupils were analyzed.

The survey concluded that

> "Pupils taking part in compensatory education reading programs were not progressing fast enough to allow them to catch up to nonparticipating pupils.
>
> A number of pupils among both participants and nonparticipants had reading achievement levels below national norms. For both participants and nonparticipants that 'deficit' grew progressively greater in each succeeding grade level samples."[13]

The survey also found evidence indicating that pupils with the greatest gains were among the less socially disadvantaged of the sample.

> "High gain pupils [came]. . . from families of higher income, their parents had more education, the occupations of the parents had greater skills, and they were predominately white.[14]

The 1968 survey found that compensatory reading programs at grades 2, 4 and 6 had virtually a random chance either of improving or worsening children's test scores when compared to similar children not recepients of Title I funding.[15]

Originally, the 1969 survey, which used essentially the same research design as the 1968 evaluation, ". . . was intended to replicate the 1968 effort" and thereby provide further evidence of compensatory education effectiveness. But following the disappointing findings of that study, it came as little surprise to USOE that a similarly designed study would yield an equally discouraging conclusion.[16] McLaughlin suggests some sort of cover up by USOE contending that the education office chose not to release the manuscript, after Gene Glass, the director of the survey, officially informed it of his negative findings.

> "Although compiled and printed, it was not 'available' in the summer of 1970 even to qualified researchers under contract to DHEW's Title I Task Force. Gene Glass himself was not, at that time, able to distribute copies of the document, and the report now remains buried somewhere in USOE. Thus the report that was to provide 'definitive' information on the efficiency of implementation of Title I ESEA and the effectiveness of that program has never seen the light of administrative day, nor has it (officially) informed a single decision-maker. Since it failed to serve its main purpose for USOE--to provide positive data for a report to Congress--the report has been for all practical purposes, suppressed."[17]

By the end of the decade, the only nationally sampled study of the cognitive impact of compensatory pre-schools was an evalution of Head Start undertaken by Ohio University and the Westinghouse Learning Corporation (Ohio-Westinghouse). Basing its analysis on a sample of 104 Head Start Centers, Ohio-Westinghouse's 1969 report found the summer Head Start programs to be only "marginally effective," but concluded that many well planned, structured full year programs were improving significantly the participants' aptitudes (children in more unstructured socialization programs were less likely). Indeed, many Head Start children who had begun the program well below advantaged children on I.Q. and reading readiness measures approached the national norms by grade one. Unlike the first annual Title I study, Ohio-Westinghouse reported greater gains for black children and for children in the Southeastern states and central cities. These encouraging reports were tempered considerably, however, by longitudinal data which indicated that the initial gains were not sustained. By the end of the second grade, most Head Start children who had made relative gains during the pre-school years were little different on cognitive measures than similar children without pre-school experience.[18]

The apparent failure of the Ohio-Westinghouse Study to produce sustained preschool gains contributed powerfully to the bewilderment felt in many educational circles at the end of the sixties. Tremendous faith had been placed in the powers of early enrichment, during the so-called "magic years," but even the politically popular Head Start could not produce a significant lasting effect. To compound the problem, a new and more conservative Nixon administration in Washington could hardly be expected to dismiss the negative findings of the early national studies and to endorse enthusiastically a continuation of compensatory education. In a 1970 speech Roger Freeman, a White House education advisor, was quoted summarizing his impressions of the evaluations of Title I programs.

> "We now spend more than $1 billion a year for educational programs under Title I of the Elementary and Secondary Education Act. Most of these have stressed the teaching of reading, but before-and-after tests suggest that only 19% of the children in each program improve their reading, significantly; 13% appear to fall behind the two-thirds of the children remain unaffected--that is they continue to fall behind."[19]

The disappointing findings of the early national evaluations may have been distorted by what may be called a "canceling effect," a concept that may be best understood if we recognize that educational evaluators have typically used the figure 1.0 to describe the mean reading of math achievement gain in a single year. Since subjects in the national studies constituted a heterogeneous group exposed to an infinite variety of teachers employing a myriad of methods, what "works" for some children may not be effective for others. Consequently, students with appropriate instruction who are really benefiting from Title I (achieving above the norm at say 1:1) were averaged with children only achieving at .4, for whom compensatory education had been a meaningless or even negative experience. Since the large scale evaluations were unable to continue for what may have the relevant demographic variables and effectively isolate particular kinds of instruction, lumping together test scores such as these may have revealed that the overall academic growth was only .7 to .8 and that Title I was ineffective. The fact that some children in the early national samples did appear to be benefiting, at least initially, from compensatory education may lend credibility to this position.

Furthermore, if the "canceling effect" hypothesis had validity, movement to smaller studies of individual programs, serving more homogeneous populations, should produce more encouraging results.

PROGRAM EVALUATIONS TO 1975

Given the many problems associated with conducting and interpreting studies with large national samples, a popular alternative strategy for assessing compensatory education effectiveness has been to search the nation for successful enrichment programs. Typically, through an extensive reading of published and in-house program evaluations, a research organization would gather information on several hundred programs. Researchers would then make site visits for further study to those programs reportedly making month-for-month or average achievement gains that appeared to be based on hard data. If the closer scrutiny confirmed that the programs were successful, the research organization would gather additional information on curriculum and methodology so these exemplary programs could serve as models for other educators constructing compensatory programs.[20]

The disappointing findings of not only Coleman and the 1968 and 1969 Title I reports but of a U.S. Commission on Civil Rights Study of pre-ESEA compensatory education programs gave even greater meaning to the USOE sponsored searches. In its report, Racial Isolation in the Public Schools, the Commission reviewed evaluations of Higher Horizons and Project Bannecker, perhaps the two most publicized local precursors of Title I, and a number of programs funded by the Ford Foundation's Great Cities Improvement Project (See Chapter Four). According to the Commission, longitudinal test scores from Higher Horizons and Project Bannecker did not support earlier optimistic findings and the data from ". . . more than 20 other compensatory education programs in large cities. . .in most instances. . .did not show significant gains in achievement."[21] By the time the Commission's findings were published in 1967, the Office of Program Planning and Evaluation personnel were conceptualizing reviews of not 20 but more like 2,000 programs. It was hoped that a nationwide study of this many programs would be able to identify, in a relatively short period of time, a number of programs which were working.[22]

The earliest, the most publicized, and eventually the most extensive search for exemplary programs was conducted by the American Institute for Research (AIR) of Palo Alto, California. Under contract to USOE, AIR identified thirty-one exemplary programs from pre-school to grade twelve in its first two reports which became the basis for the It Works Series, a thirty-one booklet package published by USOE providing detailed descriptions of each successful program.[23] Twenty-one of these programs were selected in the 1968 report after AIR reviewed written reports of over 1,000 compensatory programs in existence from 1963-68, chiefly by searching ERIC, libraries and collecting data from over 300 mail requests.[24] Using the same process eleven more programs were identified in the 1969 publication and another ten selected in a third report released in 1971.[25] In order to qualify for a site visit by AIR researchers, a program had to have some hard data indicating pupil achievement gains of at least a month's learning for a month of instruction. Programs that only matched but did not exceed the national achievement rate were labeled only "moderately successful." "Successful programs" were only those that produced gains greater than the 1:0 (average) growth rate since AIR took the position that disadvantaged children could only catch up to the national norm if they exceeded the achievement growth rate normally attained by advantaged children.[26]

In a 1972 summary report, AIR explained that the forty-one exemplary programs identified by the three studies from 124 sites visited represented only 2.3 percent of the more than 3,000 documents initially reviewed.[27] In the 1971 publication AIR lists the four primary reasons for program rejection as "(1) inadequate sample selection, (2) failure to employ reliable and valid instruments, and (3) failure to demonstrate statistically, any (4) educationally significant cognitive benefit."[28]

> This study confirmed the conclusion of the earlier two studies in this series; namely, that very few compensatory education programs for disadvantaged children have clearly demonstrated success. . .It should be pointed out that most of the programs rejected during this study were not rejected because they were demonstrated failures, but rather because their evaluation methodology was so inadequate that a conclusion about success or failure could not be drawn. Clearly, improvement must be made in program evaluation before the effectiveness of compensatory programs can be fairly assessed.[29]

The problem involving inadequate data and program failure continued to plague even the thirty-one exemplary programs identified in the first two AIR reports which had not only searched for new programs (begun after January, 1968) but included a follow up study of the original 31. According to the 1971 follow up study publication, of the twenty-seven still in operation only nine provided new "hard" data indicating that they had remained successful.*[30]

At this point it is appropriate to consider AIR's listing of features more characteristic of successful programs than unsuccessful programs.

"Pre-School Programs

- .careful planning, including statement of objectives
- .teacher training in the method of the program
- .small groups and high degree of individualization
- .instruction and materials closely relevant to the objectives

Elementary School Programs

- .academic objectives clearly stated
- .active parental involvement, particularly as motivators
- .individual attention for pupils' learning problems
- .high intensity of treatment

Secondary School Programs

- .academic objectives clearly stated
- .individualization of instruction
- .directly relevant instruction."

---

*Of these 27, one did not have any new data, 5 would not release their data, 7 presented inadequate data, and 5 had adequate data which indicated the program was no longer successful.

The characteristics identified by AIR as "most common" to all forty-one successful programs at all levels were the following:

"A. academic objectives clearly stated and/or careful planning
B. teacher training in methods of the program
C. small group or individualized instruction
D. directly relevant instruction
E. high treatment intensity
F. active parental involvement[31]

The AIR exemplary program characteristics seemed to give general support to the findings of earlier studies of Project Head Start and Project Follow Through (a federally funded program for primary school age pupils, see p.) which suggested that well-planned, structured programs produced greater achievement gains. Further support for the importance of well-defined objectives and planning, as well as AIR's finding correlating achievement and parental involvement, came from the Center for Educational Policy Research (CEPR) at Harvard. The CEPR study reviewed some 750 Title I program evaluations at the elementary school level (grades one to three) focusing on two instructional strategies (structure and parental involvement) that earlier research had suggested were effective in increasing the achievement of disadvantaged pupils. Regarding parent involvement, the study reported some evidence that parent training programs, which helped the parents learn to be effective teachers in the home, can affect achievement gains." However, "more involvement in school affairs seems not to have this result.[32] The paper on structure concluded that most compensatory programs were not highly structured, but *all* that were reported encouraging achievement gains.

> We found that highly structured, prescriptive and teacher directed programs were extremely atypical of Title I programs and thus constitute a small sub-sample of project.* However, *every such program we located reported a minimum of a month gain on standard tests of verbal ability for every month of school*. Results of the more typical general enrichment programs, on the other hand, were highly variable. While a few such programs met the minimum success standard of 1:1, most did not, even in this universe of SEA nominated 'successful' programs.33

---

*Of 672 programs sent to the Center by SEAs throughout the country as promising projects, only about 10 percent were described as structured. The Center appealed to the State Departments of Education to specifically identify additional structured programs and forty more were added to the sample.

Eighty-five percent of the evaluations reviewed by CEPR had data that was so inadequate that these programs had to be eliminated from consideration; consequently, the Center reviewed only thirty-four structured programs. Nevertheless, it was encouraging that each of these programs reported achievement gains of 1:1 or greater.[34]

The Center warned the reader, however, not to use their limited data to reach any conclusions that Title I "works" or even that structured programs represent a very small sample. In addition:

> Evaluations are done to satisfy several different groups of people--seldom is the researcher on the top of the list. Evaluations are often political documents, and must be read with that in mind.
>
> The most serious problem presented by the evaluations, however, stem from the lack of control, lack of randomization, and the concomitant possible confounding of treatment effects. Teacher differences, pupil characteristics or other programs/experiences in the school, not the treatment itself, may account for post-test gains. For example, most Title I evaluations do not make selection criteria clear. When--as is sometimes the case--children are chosen for their potential rather than degree of educational disadvantage, the likelihood of impressive gain scores increases.[35]

CEPR feels that inflated pupil potential may have been the case at one of the sites they visited in Robbindale, Minnesota, a Minneapolis "inner ring" suburb receiving Title I funds for the bottom 8 percent of underachievers in the Robbindale District. Only 750 of 30,000 Robbindale students came from welfare families and comparing underachievers in this district with low achievers in nonaffluent North Minneapolis may be like comparing "apples and oranges."[39]

CEPR offered examples of typical "structured" and "general enrichment" Title I programs which highlighted their philosophic and methodological differences. The "structured" example zeroed in on attacking reading retardation by diagnosing the child's reading problem and then building a "specific" "highly individualized" program to meet his needs. There was no room for a general haphazard approach ("remedial instruction should be organized instruction"); teachers were expected to know a pupil's ". . .expected sequence of word recognition skills," understand the student's levels of comprehension" and "keep a good cumulative account of [the] child's progress." Emphasizing a diagnostic prescriptive methodology did not deny a teacher's expression of humanity. Indeed, the guidelines for teachers spelled out the necessity of considering the "child's personal worth" and the importance of offering a student ample encouragement.[40]

CEPR contended that a "general enrichment" program usually called for "multiple program objectives reflecting attention to the 'whole child'--e.g., cognitive, affective and physical objectives." The program content was seen as ". . .often based on a general inventory of student grade level needs, rather than individual diagnosis and prescription" and as ". . .often merely an extention of typical classroom methodologies." The report offered as typical a Title I program which gave attention not only to pupils' academics ("reading, writing, spelling, listening and talking") but to the children's health ("so that they may have the physical and emotional stamina to learn to live and like to learn") and self-acceptance.[41]

While it may be correct to state that programs designed specifically to teach reading have a better chance of meeting that objective than those constructed to develop the whole child, the fact that general enrichment programs did not emphasize the formal monitoring of pupil progress could have weighed heavily against their identification by CEPR as producing 1:1 or better pupil gains. Eighty-five percent of the evaluations reviewed by CEPR had data so inadequate that their programs had to be eliminated from consideration.[42] Because the structured programs were more likely to give greater attention to keeping good records on achievement growth, it should come as no surprise that their success rate should far surpass the general enrichment projects.

Another important exemplary program search in the early 1970s was conducted by the RMC Research Corporation, but RMC reported no common characteristics of its selected programs.

> ". . .there is no single key to success in compensatory education. What characteristics make the selected project work while so many others fail can only be the subject of speculation at the present time."[43]

Under contract with USOE, RMC was asked to identify programs that were both successful (with "achievement gains at least one-third of a standard deviation greater than expectations based on national norms or control group scores") and could be packaged for possible dissemination nationwide. Originally, RMC attempted to select for dissemination eight models identified as exemplary by earlier researchers such as AIR but could find only three that met the corporation's "rigorous established criteria." Consequently, RMC had to scan over 2,000 projects before finding six* that met their criteria for achievement, cost and replication.

Despite RMC's contention that there was no "single key to success in compensatory education," a reading of the project descriptions of the six models in OE's Educational Programs that Work suggests that strong emphasis on basic skills acquisition, careful planning and intensity of treatment are characteristics common to all six programs.[44] Methodologically, all appear to prescribe specific remediation, only after a careful diagnosis of pupil deficiencies. But assuming these denominators, RMC found that educationally disadvantaged students may enjoy educationally significant achievement gains through a structured programmed approach employed by experienced teachers, through individualized tutorials by either teachers or peers or by integrating reading and math curricula with games which simulate career interests[45] (See Insert).

---

*Intensive Reading Instructional Teams, Hartford, Connecticut; Programmed Tutorial Reading, David County, Utah; Project R-3, San Jose, California; Project Catch Up of New Port Beach, California; High Intensity Tutoring Center, Highland Park, Michigan; Project Conquest of East St. Louis, Illinois.

EXAMPLES OF EXEMPLARY COMPENSATORY PROGRAMS

In terms of longevity, the programs described below may be the three most successful compensatory education programs in American history. Each one was identified as exemplary by AIR in its first report in 1968, by AIR in a follow-up study published in 1971 and by RMC in 1974. In June, 1980, these programs were still listed in Educational Programs That Work, an OE publication listing all programs approved as successful by the Joint Dissemination Review Panel.

PROJECT: IRIT: Intensive Reading Instructional Terms

TARGET AUDIENCE: Pupils in grades 3-4 who are deficient in the basic skills of reading and language. Pupils receive this intensive instruction three hours daily for one 10-week cycle. IRIT was originally approved by JDPR as a grade 3-4 program. However, the program has been used with grades 5-6 with excellent success.

DESCRIPTION: *A laboratory project for third- and fourth-grade students with difficulty in reading.*

IRIT features a carefully individualized diagnostic approach made possible by low student-teacher ratios and the use of a wide variety of materials and equipment. The unique features of IRIT are high-intensity and team teaching directed toward reading disabilities. Forty-five students per team of teachers are selected for each of the three 10-week cycles. The teachers have classes of approximately 15 pupils each and all pupils see each teacher daily. The program design includes three areas of concentration: (1) decoding, (2) individualized reading, and (3) vocabulary and comprehension. Students move from one area to the next at approximately one-hour intervals and spend the entire morning in these language arts areas. Pupils return to their sending schools in the afternoon for instruction in other basic subjects. A balanced approach to reading is provided through the three areas of instruction. The instructional design provides this content: (1) the individualized reading area provides a wide variety of reading materials and offers assignments to enrich the student's background, promote written and oral language skills, and instill pleasure in reading; (2) the vocabulary and comprehension area puts emphasis on the various sub-skills of comprehension (meaningful experiences are provided in order to promote concept development); (3) decoding, the core subject, uses an individualized approach to assist the student in his/her ability to attack new words. The specific skills-needs in phonics are taught so as to develop independent methods of word analysis. The three areas are coordinated and reinforce and supplement each other, providing a balanced reading program.

The IRIT teachers work with students for the entire morning and then spend the afternoons: (1) preparing the individualized lessons, collaborating on the skills of each student, and updating the records for each of the 45 students; (2) developing new instructions materials based on student needs; (3) meeting with teachers and parents from the sending schools; (4) participating in professional development and training sessions; and (5) providing inservice sessions for other classroom teachers. This program design prepares teachers who are very skilled in reading instruction. In addition, each IRIT Center has a team leader who is a certified reading specialist.

ASSURANCES/CLAIMS: The California Achievement Test (pre- and posttest) and the Botel Phonics Inventory are some of the tests used for evaluation. These tests are administered at the beginning and at the end of each cycle. Mean total reading scores for 1974-75 were raised from 2.6 to 3.7, a gain of one year and one month in 10 weeks. Additional test results from other years are available.

IMPLEMENTATION REQUIREMENTS: One reading specialist, the team leader, and two other teachers who have strengths in the teaching of reading. One IRIT Center requires three separate classrooms and office space for a part-time secretary and a part-time project director. Staff will need training in diagnosis and in the evaluation of test results. A school system might want to begin with one center, but a project director can easily supervise an additional center.

FINANCIAL REQUIREMENTS: Budget for one IRIT team is $59,072, serving 135 pupils in one year (approximately $400 per pupil). Start-up costs average an additional $25-$50 per pupil, depending on the specific equipment ordered.

SERVICES AVAILABLE: *An NON-funded Developer/Demonstrator project.*

Awareness materials are available. Visitors are welcome by appointment. Training may be conducted on-site and out-of-state at no cost to the adopter, paid for by the Dissemination/Diffusion Grant. Project staff is available on a limited basis to assist with training at the adopter site, also at no cost.

PROJECT: PROGRAMMED TUTORIAL READING

TARGET AUDIENCE: First graders in the bottom quartile who need help learning to read are tutored on a one-to-one basis. Originally approved by the JDRP as a first grade program, this program now operates as a first and second grade program.

DESCRIPTION: *An individualized tutoring program for first graders having difficulty learning to read.*

Programmed Tutorial Reading (PTR) supplements, rather than supplants, conventional classroom teaching. The teaching strategy employs many elements of programmed instruction: frequent and immediate feedback, specified format, and individualized pace. However, whereas programmed instruction has often sought errorless or near-errorless learning with many cues at first, followed by a fading of cues, the tutorial program proceeds in the opposite manner with minimal cueing at first, followed by increased prompting until the child can eventually make the correct responses. The 15-minute tutoring sessions are highly structured. During their sessions, children read from their regular classroom basal readers while the tutor follows exactly one of 11 tutoring programs presented in the tutoring kits. The programs in the kit specify in detail what and how to teach. They are designed so that any decision made by a tutor about a child's reading is limited to judging the correctness or appropriateness of answers. Reinforcement and praise for success are essential parts of the instructional strategy.

ASSURANCES/CLAIMS: The original evaluation of program impact was a comparison of matched pairs of students, using the Gates-MacGinitie Reading Test, Primary A, form 1. Tutored children scored at 1.9 grade equivalent at the end of first grade, while non-tutored children scored at 1.5. Later evaluations, using the same test, have produced similar results using a treatment-control group design instead of matched pairs. The area in which the program is being used is middle-income suburban.

IMPLEMENTATION REQUIREMENTS: Implementation of the project requires a staff consisting of a part-time director, tutorial supervisors, tutors, and secretarial help. Facility requirements are limited to a quiet place and a side-by-side seating arrangement at a desk or table for the tutor and the child. Basal reading materials. a tutoring kit for each tutor, supplementary reading materials, and miscellaneous materials are also required. Training requirements include extensive training of tutorial supervisors and approximately 15 hours of training for tutors.

FINANCIAL REQUIREMENTS: The annual per-pupil cost ranges from $150-$250, depending upon rates of pay of tutors and supervisors. Personnel costs, which include all administrative and clerical assistance, account for approximately 98% of total budget.

SERVICES AVAILABLE: *An NDN-funded Developer/Demonstrator project.*

Initial and secondary awareness materials as well as a Project Information Package (PIP) are available. Project staff is available to conduct awareness and pre-adoption sessions out-of-state. Visitors to the project site are welcome by appointment. Implementation-phase training is available to a limited number of adopters.

PROJECT: *PROJECT R-3, READINESS, RELEVANCY AND REINFORCEMENT*

TARGET AUDIENCE: Students of all abilities, grades 7-9.

DESCRIPTION: *A motivational basic-skill program that interrelates the reading and mathematics curricula through gaming/simulation activities involving career awareness.*

Project R-3 was jointly designed in 1967 by the San Jose Unified School District and the Education Systems Organization of Lockheed Missiles and Space Company, with the help of consultants from San Jose State University. It includes a curriculum that interrelates reading and mathematics with reinforcement through gaming/simulating intensive involvement (a three-day study trip), parental involvement, and an inservice training program for staff development. The main objective of Project R-3 is the upgrading of essential reading and mathematics skills. By deeply involving the students in classroom games and simulations, the program seeks to motivate them to achieve in learning experiences; to make them ready to learn, to make learning relevant, and to reinforce positive attitudes and behavior.

The project utilizes the diagnostic/prescriptive individualized approach in reading and math. Reinforcement of skill areas is provided through gaming/simulation activities that involve team learning, the decision-making process, and developing career awareness.

ASSURANCES/CLAIMS: The CTBS (Comprehensive Test of Basic Skills, Forms S2 and S3) shows that participants in the mathematics component (1976-77) made 2.4 months' gain for every month of instruction. For the reading component, 1.8 months' gain for every month of instruction was achieved. An outside evaluator is contracted for both product and process evaluation.

IMPLEMENTATION REQUIREMENTS: Reading and mathematics teachers should have a knowledge of the diagnostic/prescriptive approach to individualized instruction. Teachers must be receptive to team planning. All staff should develop expertise in gaming/simulation. Approximately 50 hours of inservice work is accomplished by each staff member in a given year.

FINANCIAL REQUIREMENTS: The basic materials of a reading program at the secondary level can be utilized. Special prepared math contracts cost approximately $150 for a complete set of masters to be duplicated. Other costs relate to gaming/simulation activities, contracts, and the cost of secondary instructional aides.

SERVICES AVAILABLE: *An NON-funded Developer/Demonstrator project.*

Awareness materials are available at no charge. Visitors are welcome by appointment at both project and demonstration sites. Training is conducted at the project site (adopter pays only own costs). Training may be conducted out-of-state (exemplary project staff costs may be paid for). Project staff may be able to attend out-of-state conferences (costs to be arranged).

1974 saw both the publication of the RMC models and the creation by OE of the National Diffusion Network (NDN) as the primary entity to disseminate information on exemplary programs. Discussion of such a dissemination network had been going on for some time at OE. For nearly a decade Title III of ESEA had been sponsoring a variety of locally initiated "innovative" approaches to education and encouraging the formation of "exemplary" programs for possible adoption. Once a program had been identified, however, it became increasingly apparent that a more systematic approach was necessary to disseminate information on its value. For example, the tidy It Works Series booklet, which compiled descriptions of the AIR-selected projects, may have been a useful political tool to champion compensatory education's effectiveness but its contents remained virtually unknown outside of Washington.[46]

Funding for innovative and/or exemplary programs under Title III was scheduled to terminate in 1974, this provided much of the impetus for NDN's emergence during Title III waning days. Thirty-three exemplary programs, originating from the grass roots, had been identified by 1974, and the Title III staff wanted a mechanism to assure that these model projects would be used elsewhere. Consequently, the National Diffusion Network began by offering its thirty-three programs as "Developer Demonstration Projects." NDN also began funding a State Facilitator's Project in the various states to help spread the word. Part of the adoption costs of localities wishing to capitalize on the easy availability of a research tested program was financed by NDN.[47]

Most of the thirty-three Title III programs were compensatory in nature, but several of these locally created products were in such areas as environmental and physical education.[48] To assure the dissemination of Title I programs, OE's Office of Planning, Budgeting and Evalution* soon plugged the six RMC Title I programs, which had been selected in part for their replicability, into NDN's delivery system.**

---

*Now called the Office of Evaluation.

**In 1975 the Stanford Research Institute, under contract with the Office of Planning, Budgeting and Evaluation, began a study of the infant NDN field test operations and by 1976 the data suggested that the Title III programs were being well implemented by NDN. At this point the relationship between Title I and NDN was cemented.

Indeed, by 1976 OE had contracted with CEMREO, a research organization in St. Louis, to identify an additional six exemplary Title I (and four Bilingual) programs for dissemination.[49]

Beginning also in 1974 was the Joint Dissemination Review Panel (JDRP),* a consortium of twenty-two evaluators (half from OE, half from the National Institute for Education) which passes judgment on the authenticity of programs labeled "exemplary". During the decade of the 1970s approximately 250 programs, nominated by research organizations or state educators, passed the rigorous scholarly scrutiny of JDRP and roughly 150 have been funded as developer demonstration project by NDN.[50] A scanning of the descriptions of these programs in Educational Programs That Work will suggest that many are basic skills oriented.[51] Currently, twenty-two Title I programs have been validated by JDRP.[52]

If the process of identifying programs was running smoothly by the mid-1970s, the packaging and dissemination of the selected models was undergoing growing pains. During the years 1974-76 RMC field tested each of its six exemplary programs in seventeen different schools. The programs were distributed in what was labeled Project-Information Packages (PIPs), plastic containers designed to facilitate easy distribution and minimize cost, and stuffed with adoption instructions and descriptions of the exemplary model. Although the results of the two year study found that PIPs could be rather easily implemented elsewhere, the achievement gains of children at the adoption site did not parallel the achievement growth of pupils at the original location. The Metropolitan Achievement Test (MAT) was given to participants in the adopted program for both years of the study, and the authors ". . .found few gains meeting the criterion of the one-third standard deviation above the equal percentile growth expectation."[53] However, the MAT was not used by any of the original programs and there were questions raised about its ability to properly measure the reading areas covered by the PIPs. The researchers recommended PIP's continuation and with the mixed findings in hand OE initiated yet another search for exemplary programs for PIPing in 1976.[54]

---

*Formally the Dissemination Review Panel (DRP), which originated at OE in 1972, before Congress credited the National·Institute of Education. The thirty-three Title III programs passed reviewal by DRP in 1973-74.

THE PROBLEM OF SUSTAINING EFFECTS

By the middle 1970s program evaluations of compensatory education, particularly at the pre-school level, had contributed to a growing concern about the sustenance of measured gains. The Ohio-Westinghouse national evaluation of Head Start had reported a fading of the initial I.Q. gains of many Head Start children following their exit from the enrichment experience. Now several of the experimental compensatory pre-school programs (See Chapter Four) were reporting a similar phenomenon. The description and evaluation summaries of the following projects are among those which received the greatest publicity (they were included in a 1974 OE review of pre-school compensatory education)[55] and employed the most sophisticated research design.

Bereiter-Engelmann. This behavioristic teacher-directed academic model originated at the University of Illinois and dramatically raised the I.Q.s of participating disadvantaged children from the Champaign-Urbana community by nearly twenty points. The design has been copied throughout the country, but this highly structured program with a strongly emphasized verbal component, was unable to sustain the early gains, with fade out occurring shortly after entry into the first grade.[56]

The Ypsilanti Perry Pre-School Project. This experimental program began in 1962 under the direction of David Weikart. It is a structured cognitive model relying heavily on Piagetain theory on the acquisition of intelligence. The program emphasized parental involvement with a project staff member visiting the home of each child once a week. The program successfully raised the I.Q.s of participants well above the controls but by the third grade these gains had nearly washed out.[57]

Early Training Project, Nashville, Tennessee. Begun by Susan Gray and Rupert Klaus of George Peabody College in 1961, this pre-school program employed a rather structured, cognitive "whole child" approach for three summers before school entry. During the school year staff members visited the home regularly to work with mothers in the areas of reinforcement and verbal communication. But like Bereiter-Engelmann and Weikart, the Gray-Klaus research found that promising early gains of experimental groups over controls had almost faded by the end of the fourth grade.[58]

The Howard University Pre-School. This experimental program began at Howard University in 1963 to determine if disadvantaged children enrolled in a traditional nursery similar to middle class children could obtain normal I.Q.s. This goal was accomplished during the pre-school experience but by the end of the second grade the I.Q. differences between the experimental and control groups were no longer statistically significant.[59]

The Karnes Ameliorative Pre-School Program, Urbana, Illinois. Now housed in the same building as the defunct Bereiter-Engelmann pre-school at the University of Illinois. the Merle Karnes pre-school has since 1960 had behavioral objectives and structure within a cognitively oriented whole child academic model with a heavy emphasis on parental involvement. Karnes reported that the early I.Q. difference between experimental and control group children faded almost entirely by the end of the third grade.[60]

Developmental Research Laboratory, Temple University. E. Kuno Beller, the mastermind and director of this program, is one of the only well-known early childhood experimenters to measure longitudinally the effectiveness of a traditional pre-school. Emphasizing the further development of the child's curiosity and creativity, Beller reported highly significant I.Q. differences remaining between experimental and control groups at the end of the third grade. No data was provided on academic achievement with the exception of school grades which reveal that experimental girls (but not boys) have a slight advantage over controls in grade four.[61]

Learning to Learn Program, Jacksonville, Florida. A structured cognitive pre-school project, the Learning to Learn nursery school raised the I.Q.s of participants on the average of twenty points (87.7 - 107.4) while control group children in a traditional pre-school were unaffected (88.1 - 86.8). Highly significant differences have been reported between the two groups at the end of the first grade (107.0 - 91.1) and at the end of the second grade with another wave of children (103.6 - 86.3).[62]

Initial Enrichment Program, New York, New York. This program was initiated by Martin Deutsch in 1958 and became a comprehensive five year program from pre-kindergarten through the third grade with a heavy emphasis on language development, self-discipline, and individualized instruction. At the end of the third grade one wave of experimental groups maintained a slight but significant advantage over control groups in I.Q.[63]

If one judges pre-school compensatory program effectiveness by its ability to demonstrate a sustenance of initial gains, by 1975 the Learning to Learn Program, the Developmental Research LAB and the Initial Enrichment Program appeared to be the most successful. Under contract with OE to analyze the compensatory pre-school data, Urie Bronfenbrenner did not comment on positive findings of Learning to Learn but did raise doubts about the validity of the Beller and Deutsch findings. Bronfenbrenner argued that "motivational effects" may have inflated the Beller results.

> [There]. . .is the possibility of motivational bias in favor of the nursery families who were self-selected through their positive response to a written invitation sent out by the schools and against the children in the . . .comparison group, whose parents did not enter them in school until the first grade.[64]

Regarding the Deutsch project--a five year intervention program that enrolled children at age three through the third grade--Bronfenbrenner remarks that in the case of at least one wave of children ". . .the means for the experimental group showed the characteristic hairpin turn while the children were still in the program." He stated that "at the final testing, after the children had been exposed to five years of the intervention, the I.Q. difference between the experimental and randomized control group was a non-significant four points."[65]

In 1975 Duetsch's study of his Initial Enrichment Program was one of the few published longitudinal reports which followed the same children in a post kindergarten level compensatory education program. In 1971 the American Institute of Research published a follow-up study of its twenty exemplary programs of 1968 and 1969, but it offered no clues as to why only six of the twenty-seven still in existance were no longer labeled "effective."[66]

More Effective Schools

One of the original AIR programs, New York City's More Effective Schools (MES), released an evaluation in 1969 which may explain why AIR chose not to call it exemplary two years later. According to an evalution of MES by the Psychological Corporation (PC), shortly after the program began in September, 1964, this general enrichment elementary school project, operating in seventeen schools, reported impressive achievement gains (well above 1:1) on the Metropolitan Achievement Test (MAT) in word knowledge and reading. The reported gains, however, were not sustained. According to the PC, the reading achievement scores on the MAT of children who entered the program in the second grade did not differ significantly at the beginning of grade two from the reading MAT scores of children in the control schools. By the end of the third grade, however, the average MES third grade was not only well ahead of the controls but surpassed the national norm in word knowledge and reading. In the 1968-69 school year the PC analyzed the MAT reading scores of only those MES and control group children who had taken the tests in the second and third grades, nearly four years after the original testing in the fall of 1964. At the end of the fifth grade ". . .differences between the means of the groups of paired MES and control schools in word knowledge and reading were not large enough to be statistically significant." The report explained that "the means of both groups fell below the national norm for this grade level (5.7) by three to eight months."[67]

Higher Horizons Program, New York, New York. One of the earliest and most influential compensatory education programs, Higher Horizons served as a model for Title I. A few weeks after ESEA was passed by Congress, New York City's Board of Education released an evaluation of Higher Horizons that found it to be ineffective. In 1965 the program was terminated.[68]

Table I

Differences Between Actual and Expected Mean Grade Score for Reading Comprehension of Pupils on the Three Testing Dates

	Nov., 1959	April, 1960	May, 1961
Actual grade score	2.73	3.46	4.21
Expected grade score	2.87	3.34	4.26
Difference	-0.14	+0.12	-0.05

From W. Wrightstone, *et al*'s Evaluation of Higher Horizons Program for Underprivileged Children, Bureau of Educational Research, Board of Education of the City of New York, 1964, p. 50.

The evaluation's longitudinal study on pupil reading achievement (See Table I) indicates a pattern similar to the early gain and fade out reported in the MES evaluation. The Board of Education Study of Higher Horizons included the Metropolitan Reading Test (MRT) scores of 855 third grade pupils who took the tests for the first time on November 4, 1959, less than a month after Higher Horizons was initiated. The mean grade scores for the 855 children at that sitting was 2.73 which was roughly a month and a half behind the expected grade score of 2.87 for this population. Approximately six months later, Higher Horizons children again took the MRT and the test scores of the same 855 pupils were obtained. During this period these third grade children had made a substantial gain in reading achievement. Their April grade score was 3.46 which was more than a month above the expected grade score of 3.34. At this point most Higher Horizon participants were exceeding a month's learning for month of instruction and closing the achievement gap. Roughly thirteen months later on May 16, 1961, the MRT was given once again, and the test scores of the same 855 students, now concluding the fourth grade, were analyzed. This time the mean grade score was 4.26. In this thirteen month interval most Higher Horizons children (while still enrolled in the program) had begun to fall behind.

The apparent fading of early gains that occurred in MES and Higher Horizons does not necessarily mean that the participants learned more during their first school year exposure to the program than they learned in subsequent school years. A review of multi-year state Title I evaluation reports published by the Stanford Research Institute (SRI) in early 1976 reported that ". . . the rate of gain for each grade [one to twelve] appears to be relatively constant."[72]

Because state evaluations of Title I have often been less than empirically sound, one should not place too much faith in their findings. Nevertheless, the SRI state data did show achievement gains averaging 1:1 during the school year, but in reviewing test scores of pupils from grades one through twelve, SRI reported calendar year (12 month) gains averaging no better than .7. SRI attributes part of this loss to the relative losses experienced by disadvantaged children over the summer.[70]

> "In GE scales, the achievement rate of the average student is defined to be one month-per-month during the nine month school year and one month over the summer [.9 + .1 = 1.0 for a calendar year]. An aim of Title I has been to raise the achievement rate of disadvantaged students to this month-per-month level . . .
>
> Folk wisdom assumes that the disadvantage student's achievement follows the same pattern as that of the average student but at a slower rate. This rate is usually estimated at approximately 70% of that of the average student.
>
> An extensive search of all the available empirical evidence, however, leads us to conclude that the disadvantaged student does not gain 0.7 months during the summer and therefore does not follow the pattern of the 50th-percentile student. The disadvantaged student either makes no gain during the summer or loses a month."[71]

Therefore a month for month (1:1) school year learning rate for disadvantaged pupils was not enough to prevent a cumulative deficit in achievement. With no gains in the summer or perhaps a month's loss during this period, disadvantaged pupils would have to learn at greater than month for month during the school year to keep from falling further and further behind. SRI suggested that perhaps as much as one-half of the progressive achievement gap between advantaged and disadvantaged students was caused by relative losses the latter population suffered during the summer.[72]

If one-half of the deficit could be explained by differing summer learning rates, what else contributed to maintenance of an achievement gap among Title I pupils achieving at 1:0 during the school year? SRI zeroed in on the standardized test scores from three states (California, Michigan and New York) that had a uniform testing system. If there were sustaining effects of Title I then in states reporting Title I gains of 1:0 or greater at several grade levels, ". . . we would expect to detect an upward shift in the state test results for low percentile students." Since SRI found little evidence of any "upward shift" for underachieving students, it suggested that not only during the summer ". . . but over school years in which students returned to regular classes instead of Title I."[73]

SRI concluded its report by finding it very encouraging that, for the past several years, ESEA Title I must be judged a success by the month-per-month pre- to post-test criteria, but warning that, " . . . evidence presented in this report convinces us that the ultimate success cannot be assessed in this time period."[74]

The SRI findings raised the question whether compensatory education can enjoy multi-year effectiveness without long term plans and more careful longitudinal monitoring of pupil learning. Originally, Project Follow Through, begun in 1967, was designed to sustain Head Start gains by providing compensatory continuity in the primary grades, but limited funding appears to have made it difficult for the Project and its evaluators to do more than implement and compare the effectiveness of various early childhood models.[75] Occassionally, a project site would report longitudinal data (test scores on children participating in both Head Start and Follow Through,)[76] but by 1975 the only studies of national significance were preoccupied with showing specificity of effects.*

The "discouraging" analysis concluding the SRI study was indicative of the prevailing mood of many people associated with compensatory education in 1975. Following the disappointing funding of the national evaluations and an almost frantic search for what works, compensatory policy makers began disseminating those few programs that had proven effective. But questions arose over whether exemplary program adopters could equalize the gains of the mother project and whether the gains of even the successful originals were only short term in nature. Most state evaluations were reporting impressive pupil gains, but the SEA reports were held suspect because of the serious technical flaws often associated with their data collection and analysis. The discouraging findings of compensatory education was not lost on the OE's evaluation staff. At a national research conference in 1977, Jerry Lynn Hendrickson of the Office for Planning and Evaluation offered the following summary of Title I reports released at the midway point of the 1970s.

---

*A 1971 report by Joan Bissell evaluated several experimental programs by grouping them into three general categories designated as "pre-academic" (behavioral and highly structured), "cognitive-discovery" (a structured Piagetian strategy developed by David Weikart), and the "discovery" (Bank Street-whole child approaches). Bissell reported that the three approaches were roughly equally effective in raising measured intelligence and improving school readiness for at least one year with only a slight but statistically insignificant advantage for the more structured pre-academic and cognitive discovery models.

> ". . .differences among Planned Variation approaches in both Head Start and Follow Through suggest a specificity of effects, such that in programs with specific objectives and well-formulated strategies to achieve these objectives somewhat more growth is found . . .than in whole child programs."

Another Follow Through evalution was released in 1971 by SRI, which compared three approaches to early childhood education that roughly matched the categories of

(Continued on next page)

> "Up until 1975 reports on the effectiveness of Title I as a whole were uniformly pessimistic. Aggregation of data in the States' Annual Evaluation Reports . . . always highlighted problems with the data and the finding that Title I children were falling farther and farther behind their peers, although they generally found month-for-month or better gains were mentioned. National evaluations also showed, for the most part, Title I children losing ground, while non-Title I children gained ground."[77]

1976 AND BEYOND

About the year 1976 began what may be a new phase in compensatory education evaluation. During the last half of the 1970s and early 1980s we find not only the appearance of evaluations with more positive findings but the emergence of a more sophisticated and perhaps more realistic means of assessing Title I effectiveness.

The New Reporting System

In 1976 USOE introduced to the states a new measurement tool entitled Normal Curve Equivalents (NCEs), a statistical measure which was scheduled to replace grade equivalents as the device to assess Title I pupil gains.[78] Although grade equivalents are an easily understood unit of measurement, they often give a gross inaccurate estimate of a child's achievement. According to the RMC Research Corporation (primary developer of not only Title I PIPs but Title I's revamped evaluation system), "Errors of several months are not uncommon and are large enough to make an unsuccessful project appear successful or a successful project appear unsuccessful."[79] One of the major problems appears to lie with test publishers who often base their grade equivalent scores (i.e. the raw score 78 equals a GES of 5.8) on "an estimate [of] testing done in some other month . . . [and] may not have actually tested any norm-group students

---

Bissell: (1) highly structured-behavioral, (2) cognitive discovery, and (3) open-"pragmatic." The trend suggesting a specificity of effects in the more structured models of the Bissell study reached statistical significance in the SRI report. SRI concluded that the sponsor groups whose approach is ". . .most structured and concentrates most explicitly on developing academic and preacademic skills showed a consistently higher level and rate of achievement measured by the pupil achievement test battery."

in the eighth month of the fifth grade."[80] Another is with children who score either significantly below or above grade level and in figuring out what skills they actually have relative to their peers.[81] In a 1976 technical paper prepared specifically for the LEAs, the RMC Research Corporation explained the phenomenon.

> ". . .consider the meaning of a below-grade-level score; let us say an eighth grader receiving a score of 5.8 in math. This is an even more complicated situation. Can we say that this student has the math skills of an average fifth grader, whatever those may be? Are all of the sixth-, seventh-, and eighth-grade skills missing? We simply do not know.
>
> Eighth graders normally take a higher level of a test, and the older but slower student may have answered some of the higher level items correctly, items that the fifth grader had never studied and were not even on the fifth-grade test. At the same time, the eighth grader probably missed many fifth-grade level items that an average fifth grader would get correct. Thus, while, a GES of 5.8 for a fifth grader means average performance of fifth-grade skills, a GES of 5.8 for an eighth grader may mean a lower level of skill over a wider range of subject matter. In any case, the eighth grader cannot be compared to an average fifth grader.
>
> A comparable problem occurs in understanding scores of above-grade level students. For example, on one widely used standardized test, a first grader who got a perfect score on the math subtest would receive a GES of 5.1. The test manual points out that this does not mean the first grader can do fifth-grade math. It means that the median fifth grader would get a perfect score on the first-grade test if any fifth graders took the first-grade test."[82]

NCEs have been billed as a vast improvement over grade equivalents. They are placed at equal intervals along a normal curve, and thus "a gain of five NCEs represents exactly the same amount of improvement for pupils at the extreme low end of the achievement distribution as it does for average achievers."[83] NCEs are similar to percentiles in that both an NCE and percentile of 50 are average and the numbers 1 and 99 represent the extreme low and high. Percentiles, however, are not uniformly intervaled and thus mean different things at various points in the distribution. For example, a three percentile point gain is much greater at the extreme ends of the normal curve than it is in the center[84] (See Figure 2).

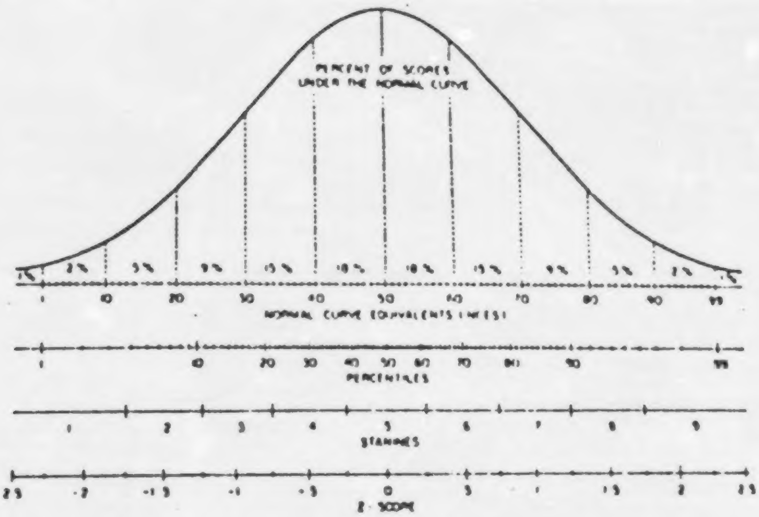


Figure 2 From G. Tallmadge: *Interpreting NCEs* Mountain view, California: RMC Research Corporation, October, 1976, p. 3.

Technical language aside, perhaps the most valuable feature of NCE reporting is that it rids the Title I evaluator of having to compare Title I pupil growth with the gains of average children. No longer is program success contingent upon 1:1 or better gains but simply on NCE gain greater than zero. Thus a NCE increase of one or better would mean the program had a better than expected impact. Since NCE gains take on different meanings at different grade levels and since converting such increases into grade equivalent improvement is of questionable value anyway, reporting in NCEs makes it more difficult to determine whether compensatory education is either producing gains which parallel or exceed the growth rate of the average child. In a second 1976 "Technical Paper" designed for LEA consumption, RMC made success sound easy to attain and comprehend and advises against establishing any absolute criteria for selecting a model program.

> All NCE gains greater than zero are good!
> Whenever the evaluation shows an NCE gain greater than zero, it means that the Title I pupils profitted from participating in the project. In general, the larger the NCE gain, the more effective the project. It is not possible, however, to designate any specific NCE gain as the criterion for exemplary or outstanding projects. A cost-effectiveness criterion seems more appropriate. Assuming that the same number of dollars were spent, for example, a 4-NCE gain produced in a treatment group of 200 pupils might be considered as good as an 8-NCE gain produced in a treatment group of 100 pupils.[85]

Although one can find little evidence in the public record to support this assumption, it is suggested here that one factor contributing to the OE decision to substitute NCEs for grade equivalents in Title I reporting was the realization that compensatory education was unlikely to have true compensatory impact (fully substituting for out-of-school influences) and that the NCE system permitted a more sensitive measure of what little impact Title I was likely to have. If future research showed that compensatory education eliminating the progressive achievement gap, it would be all well and good. In the meantime, however, programs producing better than expected gains needed recognition as "good" ones and an improvement of one or two hundreths using NCEs would be easier to attain than one or two tenths (say from .7 to .8 or .9) using grade equivalents. Evidence in support of at least a discussion of improving the measurement of more modest expected gains is provided in a "working paper" by G. Kusten Tallmadge and Barbara Fagan of RMC. The manuscript was circulated at the highest administrative levels in the Office of Planning, Budgeting and Evaluation (OPOE).[86] Entitled "Cognitive Growth and Growth Expectations in Reading and Mathematics" the paper discusses the ". . .frequently encountered, high expectations for compensatory, remedial, and other special instructional programs." Referring to the " 'normal' growth patterns of children at different achievement levels" Tallmadge and Fagan make the following observation:

> ". . .bringing underachieving groups up to grade level, or raising their performance by a quartile (or even by a stanine), would require instructional efforts two, three, or even five times as effective as the combination of maturation, regular school programs, and all out-of-school learning experiences. Such expectations appear generally unrealistic, and are more so when encountered in conjunction with projects or programs that provide only 15 to 30 minutes of instruction per day.
>
> Unrealistically high expectations will usually go unmet, thus providing a source of discouragement to educators and educational administrators alike. A more serious problem, however, stems from the fact that the expectation of large treatment effects is likely to lead to evaluation designs that are insensitive to whatever real impact the special instruction may have. For these two reasons, it seems important to analyze the kind of growth that can be expected under normal conditions and to use the resulting information as a framework against which to set objectives for educational interventions and to evaluate their effectiveness."[87]

Recent Major Evaluations

At this point it would be convenient to conclude the review of the research on Title I by discussing several recent empirically sound national studies which reported only modest NCE's achievement gains for Title I pupils. One could then summarize the findings as suggesting that our largest compensatory education program is indeed improving pupil achievement but that it may have been naive (an "unrealistically high expectation") to assume that the schools could completely compensate for "all out-of-school learning experiences." Unfortunately, the muddled world of educational research has never permitted simplistic conclusions about schooling effectiveness, and the findings of the three major studies of Title I since 1975 allow no exceptions to this dictum in the field of compensatory education. However, the Compensatory Reading Study, the Compensatory Education Study and the Sustaining Effects Study may provide the strongest evidence to date that Title I might be "working" and at least one of the reports may give us some idea of how well and for how long.

Released in 1976, the Compensatory Reading Study tested over 50,000 compensatory education (CE) and non-compensatory education (NCE) students in grades two, four and six during the 1972-73 school year, "a sample that is close to being truly nationally representative. . ." for that age group. Generally speaking, achievement gains in reading of CE students compared favorably with NCE pupils. According to the Technical Summary:

> "The nature of the results tended to vary somewhat depending upon the subtest and grade level being examined and upon the analytic technique employed. All of the analyses supported the assertion that compensatory students tend not to fall further behind non-compensatory students during the academic year (the main exception was for a scale called word knowledge)."[88]

Several serious questions concerning the study's assessment measures, however, marred the positive findings. According to George Mayeske, OE sponsor of the study, there were many significant technical problems, among the most serious being "lack of proper control for statistical regression."[89] Conducted before the NCE reporting system became fashionable, the Compensatory Reading Study reported pupil achievement gains by listing the average fall and spring Metropolitan Achievement Test raw score totals for each grade tested.[90]

In 1977 the National Institute of Education released the first of two reports on the effects of Title I services on pupil achievement. The "Instructional Dimensions Study" (conducted as part of the Compensatory Education Study) provided test data on pupils enrolled in fourteen projects scattered throughout the nation; each project had been nominated by various localities and state education department for their merit. In the 1977 report NIE found impressive academic year gains in the first and third grade Title I reading and math pupils participating in the study.[91]

> "In general, the results were encouraging about the effectiveness of compensatory education. First graders made average gains of 12 months or 12 percentile points in reading and 11 or 14 percentile in mathematics. Third graders made average gains of 7 months or 9 percentile points in reading and 12 months or 17 percentile points in mathematics."92

Perhaps even more encouraging were the results of NIE's "Instructional Dimensions

Follow-Up Study," housed as a chapter within a final report the Institute made to Congress in 1978 on the Compensatory Education Study. In the follow-up study a representative subsample of CE pupils in seven of the fourteen original districts were retested the following fall and their scores compared with those of a sample of non-CE pupils (scoring at the 50 percentile or above). Contrary to the findings of other similar assessments of summer learning, the Instructional Dimensions Follow-Up Study reported ". . .that compensatory education students can maintain impressive gains over the calendar year" (fall to fall) which "raise questions about the generality of summer loss for compensatory education students." The second study found no "loss in absolute terms" (achievement decline) over the summer and no "consistent or sustaintial relative loss" (learning inferior to non-CE pupils during this period."[93] (See Table II) According to the report,

> "The data indicate that low achievers do lose some ground to higher achievers over the summer period. However, the size of the differences varies considerably by grade level and subject matter, and only for first grade reading is the difference statistically significant."[94]

The study also measured the summer growth of a small percentage of the original CE sample that attended summer school and did not find summer school participants gaining greater than non-participants between the spring and fall testing.[95]

One finding of some concern, however, would appear to be the fall gains of third grade CE pupils in reading. Their modest fall to spring gains combined with their spring to fall gains gave them a twelve month increase of only .8. This compares with a 1:6 gain of their non-CE counterparts during the calendar year. (See Table II)

The third major report on compensatory education to appear since 1975 is the "Sustaining Effects Study." Because of the growing concern about the durability of gains, the Sustaining Effects Study promises to be one of the most important research projects ever conducted in the field of compensatory education. The contract for the study was awarded by OE to the System Development Corporation in 1975. Reading and math achievement test scores were collected from national samples of elementary (grades one to six) CE and non-CE pupils for three consecutive years beginning in 1976.[96] Unfortunately, at press time only test scores for the first calendar year have been

TABLE II

MEAN ACHIEVEMENT GAIN SCORES OF NON-CE STUDENTS WHOSE PRETEST SCORES WERE AT OR ABOVE NATIONAL NORMS AND OF CE STUDENTS WHOSE PRETEST SCORES WERE BELOW NATIONAL NORMS

Study Group	Relationship to National Norm	n	Fall-to Spring Gain	Spring-to-Fall Gain	Fall-to-Fall Gain
Grade 1 reading	At or above (non-CE)	296	56* 0.5**	10*** 0.1	66 0.6
	Below (CE)	344	69 1.2	0*** 0.0	69 1.2
Grade 1 math	At or above (non-CE)	435	39 0.5	6 0.1	45 0.6
	Below (CE)	97	43 1.1	2 0.1	45 1.2
Grade 3 reading	At or above (non-CE)	305	36 1.0	21 0.6	57 1.6
	Below (CE)	512	44 0.5	8 0.3	52 0.8
Grade 3 math	At or above (non-CE)	178	62 1.6	7 0.2	69 1.8
	Below (CE)	306	64 1.2	-1 -0.1	63 1.1

---

*Expanded Standard Scores

**Grade-equivalent scores

***Difference in gain significant (P .01)

From Compensatory Education Study, Final Report to Congress from the National Institute of Education, September, 1978. The National Institute of Education, U.S. Department of Health, Education, and Welfare, Washington, D.C., p. 86.

analyzed, but the limited data have already yielded some interesting information. During the academic year CE students (Title I pupils and participants in locally funded compensatory education programs) do not equal the gains of the average student but usually exceed the growth rate of pupil samples in need of, but not receiving, compensatory education.[97] In the various papers reviewing the preliminary findings grade equivalent scores are not employed to demonstrate gains (and for some reason nor are NCE's) but comparisons are made by verticle scale scores, a graphic technique which can demonstrate the growth rate of various samples. Growth patterns are also recorded in percentiles. A glance at Figure 3 reveals that pupils in need of, but not enrolled in, CE had virtually a random chance of increasing or decreasing their Fall percentile score during the academic year. On the other hand, Title I students demonstrate an upward trend at every grade level. One problem area, however, appears to be the reading achievement growth rate of CE pupils in grades four to six. Although these students made modest percentile gains, they appeared to be doing no better than their non-CE comparisons (See Figure 3). Similar to the NIE follow-up study, the Sustaining Effects Study found no absolute loss for either regular or CE students between Spring and Fall but "a small loss over the Summer for CE relative to non-CE students." Once again, summer school was found to have little or no effect on reading and math achievement scores.[98]

In the area of pre-school compensatory education, the remarkable perserverance of many of the originators of the experimental programs of the early sixties has finally begun to pay robust dividends. Following the disappointing evalutions generally showing a near washout of the initial pre-school I.Q. gains by the middle elementary grades, twelve of the pioneer early childhood investigators formed the Consortium for Longitudinal

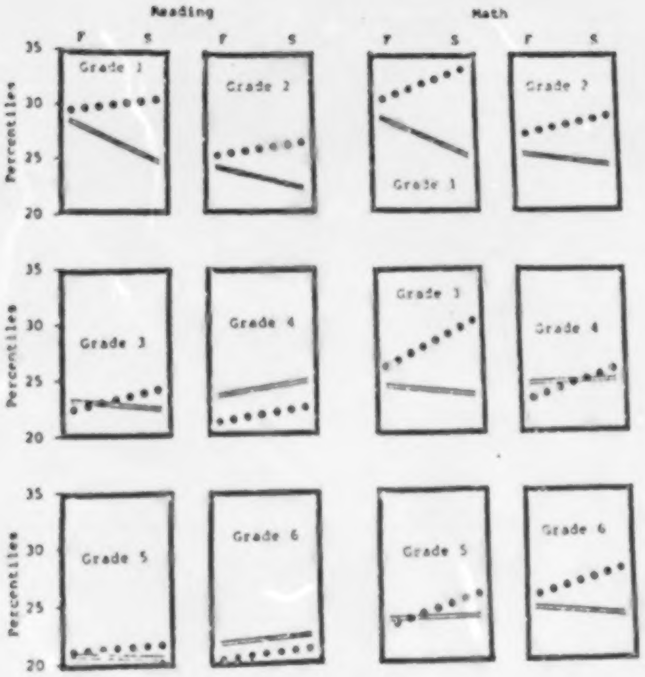


Figure 1. Reading and Math Gains, Fall to Spring in Percentiles, for Title I Selectees (•••••••) and Non-CE Selectees at Non-CE Schools Judged by Their Teachers to be in Need of CE (——). From Ming-mei Wang, *Evaluating the Effectiveness of Compensatory Education*. Paper presented at the annual meeting of the American Educational Research Corporation, Boston, Massachusetts, April 7-11, 1980.

Studies. According to Richard Darlington and other coordinators[*] of the follow-up work, the consortium "pooled their original data and conducted a common follow-up study in 1976-77 of the original participants and control groups." (Additional pooling and follow-up analyses by consortium members have continued unabated into the early 1980s.) Darlington, et al. inform us that although pre-school experiences have apparently had no sustaining effects in measured intelligence (reporting an I.Q. ". . .tapering down to smaller but statistically significant effects three or four years after pre-school and vanishing thereafter")[**], it did appear to have a profound influence on "whether a child had been retained in grade (held back) or assigned to special education classes at any time during his or her school career." 1,599 of the roughly 2,700 children in the eleven original consortium projects were available for the 1976-77 follow-up and appropriate data existed on 802 subjects in seven projects. The data shows that experimental subjects were nearly twice as likely as controls (44.7 vs. 24.1) either to repeat a grade or to be recommended for special education. Apparently the pre-school compensatory experience can provide the foundation for environmental change which affects behavior many years later.

Despite the failure of the pre-school programs to modify measured intelligence permanently, one cannot dismiss the multiyear influences of early childhood education as being exclusively affective in nature. In most cases the I.Q.s of compensatory pre-school participants were significantly higher than non-participants three or four years later, and an advantage in math achievement for the experimental subjects has been reported lasting throughout most of elementary schooling (.5 grade equivalents favoring pre-school children in the fourth grade.)[99] According to Darlington, et al., ". . .if one year of enrichment still has effects three or four years later, then twelve years of enriched schooling might have lifelong effects."

The Consortium for Longitudinal Studies and the Sustaining Effects Study have not been the only attempts since 1975 to measure compensatory education longitudinally. At least two state departments of education have monitored the achievement of the same Title I pupils for two consecutive years and both suggest a decreased growth rate during the second year.

---

[*] Darlington and Irvin Lazar of Cornell University, who did not conduct pre-school programs, have been the principal coordinators. Jacqueline Royce, Ann Sniper and Harry Murray, also of Cornell, have assisted Darlington in preparing at least one written summary of the consortium's major activities and findings.

[**] In Chart one will note I.Q. data was not given for Gordon's research which included the once promising Learning to Learn Program.

A 1979 report from Rhode Island measured the two year growth of Title I children who were in the second and fourth grades at the initial testing. Pupils who remained in Title I for the two year duration of the study comprised the "two-year" group; students who for some reason left Title I after the first year of testing constituted the "one-year" group. (See Table III). The achievement data on the "one-year" group should be of particular interest, for it represents perhaps the first reported assessment of what may happen scholastically to Title I children after exiting the program. Apparently most children in the "one-year" group graduated to the mainstream because of the success they attained in Title I. According to the Rhode Island Study, "While there were generally small insignificant differences [between groups one and two] in reading achievement at pre-test time, by the end of the 1974-75 school year [the first year of the study] the one-year group scored significantly higher than the two-year group on most measures."

Table III includes the NCE scores for the two groups of second and fourth graders for two academic years. It is evident that both groups at both grade levels made larger gains the first year and that there was actually a NCE decline during year two. The grest losses were suffered by group one during the year after their termination from the program. Despite the second year decline, however, growth ". . .across a two year period [still] . . . exceeded what would have been expected if students had not received compensatory education instruction." The study's authors suggested that the second year decline was caused in part by the often reported "dramatic" drops in achievement of disadvantaged children during the summer months. This suggestion was highly speculative because the absence of a fall testing in the second year precluded measurement of summer losses.

Table III

Mean NCEs at Pretest, First and Follow-Up Posttest for Grades 2 and 4
Composite Reading

GRADE	GROUPS	PRETEST	FIRST POSTTEST	FOLLOW-UP POSTTEST
2	Two Year	28.2	43.0	37.7
	One Year	28.2	48.9	40.7
4	Two Year	21.8	35.8	30.3
	One Year	24.7	43.0	30.7

From P. De Vito and S. Rubenstein's "A Follow-Up Study of Title I Participants in Rhode Island," paper presented at the annual meeting of the American Educational Research Association, San Francisco, California, Aprii 8-13, 1979, p. 7.

A similar study was conducted by the Tucson Unified School District. Released in 1980, this report summarized a two year investigation of Title I's sustaining effects on junior high school pupils between the fall of 1977 and the spring of 1979. Table IV gives the grade equivalent, NCE and expanded standard score gains for seventh, eighth, and ninth grade pupils at four testing times (fall and spring for two years) and does not show dramatic summer losses. The Title I sample may have lost ground in the summer relative to the average child but at least in terms of grade equivalents suffered no absolute losses. (Converted to NCEs, small absolute losses are evident for the eighth and ninth grade sample during the summer, suggesting the greater sensitivity of NCE reporting to achievement differences.) At each grade level, declines in achievement growth took place during the second year of the testing. The Tucson Study's author, Stephen Powers, made the following commentary:

> "One might hypothesize that a regression effect caused the gains the first year. However, students were selected for Title I based on scores on a standardized test administered two years before the fall of 1977.
>
> These results suggest weaknesses in evaluation practices when a group of students is tested fall to spring without any follow-up study. Although students display an initial surge of achievement the first year, that rate of achievement may not be sustained the following year." 100

Unlike the Rhode Island children, however, Tucson subjects lost little ground the second year. The NCE gains for first year seventh and eighth graders were respectively strengthened and generally maintained at the end of the second year. The small NCE loss of the eighth graders and modest lost of the ninth graders were caused largely by slight NCE declines during the summer. It should be also encouraging that despite the second year decline the seventh and ninth grade participants still gained eighteen months over the eighteen month testing period. At the ninth grade level the participants exceeded 1:1 for two consecutive academic years. And thanks to the new reporting system less than 1:0 grade equivalent gain rates for second year eighth and seventh graders became transferred as infinitestimal to modest NCE gains.

Table IV

Table 2. Comparison of Gains of the First and Second Years* as Measured in GEs, NCEs, and Expanded Standard Scores

		1977-78			1978-79		
Grade	Scale	Fall	Spring	Diff	Fall	Spring	Diff
7	GE	3.4	4.3	.9	4.7	5.2	.5
	NCE	21.6	27.2	5.6	28.0	29.3	1.3
	SS	365.9	405.4	39.5	419.9	439.6	19.7
8	GE	3.9	4.9	1.0	4.9	5.3	.4
	NCE	22.0	26.8	4.6	25.6	25.7	3.5
	SS	390.2	428.6	38.4	427.0	442.9	15.9
9	GE	4.3	5.3	1.0	5.4	6.1	.7
	NCE	21.2	25.4	4.2	23.7	23.5	-.2
	SS	407.1	441.9	34.8	445.6	471.8	26.2

* Seventh Grade N = 138
Eighth Grade N = 106
Ninth Grade N = 65

From Stephen Powers' "A Comparison of Cognitive Gains of Seventh, Eighth and Ninth Grade Students," paper presented at the annual meeting of the American Educational Research Association, Boston, Massachusetts, April 7-11, 1980, p. 6.

With the exception of the concerns over the sustaining effects of compensatory education at the elementary and secondary level, the research findings should place compensatory education in a more favorable light in the early 1980s than it enjoyed in the mid-1970s. The period since 1975 has also witnessed the growth of the National Diffusion Network. Many local evaluations suggest a number of compensatory programs are being successfully adopted.[101] In JDRP terms "success" now does not necessarily mean solid evidence of 1:1 or better gains but empirically sound research suggesting "better than expected" achievement gains.[102]

Other relevant post 1975 educational research reports, not strictly compensatory in nature, have challenged what the Education Times calls the "prevailing pessimism about the possibilities of education.[103] In 1977 the National Assessment of Educational Progress (NAPE) released the results of two national surveys of reading achievement, which compared the reading achievement of seventeen-year olds, drawn from similar samples, in 1971 and 1975. On NAPE administered reading achievement tests, seventeen-year olds from low income urban areas gained five percentage points while middle class children showed no gains. Broken down racially, black children out gained white children five percentage points to two.[104] A 1979 study from England entitled 15,000 Hours attributed the schools ethos (i.e. high expectations, student feeling of willingness to consult with teachers, consistency of values) to much of the variation in achievement among London inner city secondary schools. In this country, the research of Ronald Edmund in Detroit has led to the identification of several differences between relatively high achieving and low achieving schools with predominately black, low income student bodies.

According to Edmunds,

> ". . . (a) They have strong administrative leadership without which the disparate elements of good schooling can neither be brought together nor kept together; (b) Schools that are instructionally effective for poor children have a climate of expectation in which no children are permitted to fall below minimum but efficacious levels of achievement; (c) The school's atmosphere is orderly without being rigid, quiet without being oppressive, and generally conducive to the instructional business at hand; (d) Effective schools get that way partly by making it clear that pupil acquisition of basic school skills takes precedence over all other school activities; (e) When necessary, school energy and resources can be diverted from other business in furtherance of the fundamental objectives; and (f) There must be some means by which pupil progress can be frequently monitored." (Edmunds, 1979, p.22)

As the newly appointed Assistant Chancellor for New York City Schools, Edmunds is currently both testing many of his Detroit findings and attempting to initiate more effective educational practices in New York.

In light of the more optimistic findings since 1975, is it possible to say that compensatory education is now "working," i.e., that twenty odd years of research in this field has contributed to instructional behaviors that are finally yielding results? Unfortunately, one cannot make that claim with any certainty, for more sophisticated measurement, rather than improved practices, may account for greater measured impact of programs such as Title I. Differences in research design can apparently cause significant differences in the findings. For example, when George Mayeske, Supervisor of Evaluation Specialists, at OE, was asked why the Sustaining Effects Study did not show as much program impact as the Compensatory Reading Study, he said the former investigation "collected better data." He continued that "the Sustaining Effects Study used better comparison groups to rule out certain regression effects." According to Mayeske, it is the test themselves, rather than the ineffectiveness of compensatory education, which has contributed to much of the disappointments associated with Title I evaluations:

> "We will never get any effect with the tests we use. Essentially they are the same off-the-shelf stuff we've been using for fifty years as a sorting device. Our tests should be tied to a psycho-linguistic framework involving a hierarchy of skills. We can then relate assessment devices to points in that hierarchy."[106]

Instead of data suggesting compensatory education is equalling the national growth rate (in the case of the Compensatory Education Study), apparently we now have more refined empirical inquiries (such as the Sustaining Effects Study) producing data suggesting more modest program effects or perhaps compensatory education was "working" all along, and the successes simply were not recognized. Furthermore, it may be that additional research will show compensatory education making far more than marginal impact and thus reveal it as having true compensatory powers.

Despite the flaws in research methodology and the standardized measures, which mitigate against making any "conclusions" about compensatory education's effectiveness, we summarize the findings by listing the following hypotheses, emergent from more than two decades of research on compensatory education:

(1) Title I and associated programs are contributing positively to at least short term pupil gains in reading and math. Whereas the early studies yielded disappointing results, the major research since 1975 has been virtually unanimous in suggesting compensatory education is making a difference during the pre- and post-testing period. The Sustaining Effects Study suggests this may be particularly true in math.

(2) Longitudinal evaluations of compensatory education may raise serious questions about a program's ability to substitute entirely over a sustained period for out-of-school determinants of substandard achievement. In the Tucson study, for example, children in grades seven and nine averaged month-for-month learning for a two academic year eighteen-month period (Fall of 1977 - Spring 1979) but the significant declines during the second academic year (Fall 1978 to Spring 1979) of the study, must raise the question of whether following the same children for an additional eighteen months would find them maintaining the average growth rate of the eighteen month testing period.

In discussing his findings, Stephen Powers, the author of the Tucson Study, makes the following commentary:

> "These results suggest weaknesses in evaluation practices when a group of students is tested fall to spring without any follow-up study. Although students display an initial surge of achievement the first year, that rate of achievement may not be sustained the following year." 107

What may be advisable as well are Title I assessments which follow the same children for longer than a two year period.

(3) Relative losses during the summer account for part of the cumulative deficit in achievement. Although there is contradictory data on how much decline in the achievement growth takes place during the summer, the research strongly suggests that most disadvantaged pupils experience a greater decrease in the achievement rate between spring and fall than average achieving children.

(4) Highly structured, prescriptive programs are more likely to produce greater short term gains than "whole child" general enrichment programs. Much of the research on compensatory education programs has suggested that such variables as careful planning and specificity of objectives correlates with greater pupil learning. This data would appear to be in keeping with Barak Rosenshine's research on teaching which, has reported that "direct instruction" and ". . .high levels of student engagement within teacher-directed classrooms using sequenced, structured materials" is a more effective teaching strategy than less structured approaches.[108] The compensatory education research also seems to be consistant with the findings of the evaluations of "mastery learning" or Learning for Mastery (LFM), a structured instructional sequence originated by Benjamin Bloom in the late 1960s.[109] LFM advocates that an instructional package that includes behavioral (measurable) instructional objectives, direct attention to the defined behavior, small "bite size" steps and immediate positive feedback will maximize the chances for greater pupil learning.[110] A recent analysis of the research on "mastery learning" has concluded that the "mastery" method is much superior to the

conventional techniques.[111]

One must interpret the results of the instructional effective studies with caution, however. It must be remembered that most educational research is not longitudinal in nature, and it is conceivable that multi-year reports on achievement may yield greater effects for the general enrichment approach. The long term superiority of Progressive Education over traditional was evidenced in the 1930s when the Eight Year Study found that children who had attended more progressive high schools exceeded, in grade point average, their matched comparison counterparts in every subject but foreign language.[112] Of course, the participants in the Eight Year Study were for more advantaged than the average Title I sample and different educational methods may have different effects on different populations. When the desired academic skills are not strongly associated with the childrens socio-cultural milieux (such as a foreign language for advantaged college students), they may require more formal instructional strategies to successfully acquire these skills.

(5) *Pre-school compensatory education has had little or no permanent effect on cognitive and achievement but may have had a robust influence on affective behavior.* With the exception of a slight advantage pre-school participants enjoy over non-participants in math (at least through grade four), there seems to be no permanent cognitive effects of early childhood compensatory programs. Yet the recent findings of the Consortium for Longitudinal Studies should offer considerable hope to pre-school educators. The degree of social and educational adjustment disadvantaged children attain may be strongly related to educational intervention during the pre-school years. There appears to be a growing recognition in some educational circles that if compensatory education is to maximize its chances of being effective it should be more than a half hour to an hour pullout program to supplement regular school activities. In the domain of the basic skills, disadvantaged children may benefit a good deal more by educators presenting a united front. Currently, the national Title I office is emphasizing "coordination of efforts" which will hopefully lead to the program becoming intricately interwoven

with the "regular" school curriculum.[113] If the research suggests that specificity of objectives and careful planning reap greater benefits within a single classroom, one may assume that a structured comprehensive commitment to basic skills acquisition within a given school will produce greater equality of education achievement.

Recognizing the possible potential such as coordinated approach, Pueblo, Colorado, a city with a large minority school age population, is integrating its Title I program with the district wide Maintenance and Improvement Plan for Instruction (MIP), the latter described as a "strong, well organized instructional program" . . . which ". . .provides teachers with assistance with the important classroom management tasks of setting instructional priorities. . ., identifying basic skills needs. . ., and aiding in instructing, monitoring, evaluating and reporting a student's progress."[114] Under the MIP, the Pueblo Central Administration assists school building professionals in establishing specific school wide goals to improve reading and math instruction. Typically, the selected pupil competencies professionals chose to cultivate are skills measured by the California Achievement Test, the districts major standardized assessment instrument. The city's Title I program plays an important role in the MIP's implementation by fostering close communication with principals and regular classroom teachers.[115] The Title I teachers develop and incorporate their "Title I Plan" within a prescriptive "Building Plan" which according to Peggy Seavy, Title I Director, "promotes a constant circle of implementation" and a "continual exchange of information concerning the development of each individual child."[116] Since the implementation of the Maintenance and Improvement Plan three years ago, Pueblo has reported significant district wide achievement gains. The coordinated efforts seem to be particularly successful in the predominately Chicano Title I elementary schools where pupils have equaled or surpassed the national average on the California Achievement Test for two consecutive years.[117] Chuck Sekera, Director of the Pueblo's Maintenance and Improvement Plan, attributes its success to a "much greater focus of the entire instructional program on basic skills." Continues Sekera,

> "We found that the competencies assessed by standardized tests were sufficient to constitute a core for our basic skills program. We then wrote the appropriate objectives, embraced them and found that when instruction is focused toward those objectives, measurable performance increased. Regarding Title I, the clarification of regular program goals makes it possible for Title I and other support programs to be much more effective in their efforts to compliment the established program." 118

The fact that a federally funded program such as Title I can enjoy such harmony with a local school district may be indicative of a new spirit of cooperation among educators admist a growing optimism about the effects of education. Indeed, Robert Kennedy's insistence on an evaluation requirement for Title I may have been instrumental in promoting professional partnerships and renewed hope, for Title I has been required to take a close look at itself, to come to grips with what "works" and to broadly disseminate this information. Hopefully, Title I evaluation processes and the findings of many Title I evaluations have not only been a valuable informational exchange for compensatory educators but have influenced other educators by demonstrating the significance of empirical inquiry, by identifying effective strategies and by promoting the sharing of such knowledge. Larry Bussey, one of the principal developers of the National Diffusion Network, stated it succinctly when he was asked to reflect on the value of NDN.

> "It's really too bad we didn't have NDN in 1968. If it had existed, the expertise of what constitutes effectiveness could have been provided to other programs. The question of 'is it making a difference' is very important. The process forces us to ask questions about what it is that is making a difference or no difference. The information can then be shared with others." 119

In conclusion, compensatory education has gone from a period of idealistic promise, to an era of indignation and pessimism, to a time of perhaps more realistic assessment and expectations more encouraging findings and renewed optimism. The research has identified methods and programs which make a difference, but we are still uncertain how much impact schooling can have and for how long its effect can be measured. We must also be cautious about generalizing the effects of compensatory education on educationally disadvantaged children in general to black participants in particular because the major evaluations rarely assess pupil progress by race or ethnicity. In 1980 blacks constituted only 26% of Title I enrollment nationwide.[120] From the data we have reviewed one would assume that Title I and associated programs are benefiting black and other underachievers academically but more confident statements on compensatory education's ethnic and racial influences await the findings of future research.

BIBLIOGRAPHY

Part I

1. Martin Deutsch popularized the use of the term "cumulative deficit" in his book *The Disadvantaged Child*, published in 1967, New York: Basic Books. For a concise review of the origin of the term see A. Jensen's "Cumulative Deficit: A Testable Hypothesis?," *Developmental Psychology*, 1974, v10, n6, pp. 996-997.

2. D. Hawkridge, *et al*., *A Study of Selected Exemplary Programs for the Education of Disadvantaged Children*, Part 1, (Arlington, VA.: ERIC Document Reproduction Service, ED 023 776, 1968.)

   M. Wargo, *et al*., *Further Examination of Exemplary Programs for Educating Disadvantaged Children*, (Arlington, VA.: ERIC Document Reproduction Service, ED 055 128, 1971.)

   M. Wargo, *et al*., ESEA Title I: *A Reanalysis and Synthesis of Evaluation Data from Fiscal Year 1965 to 1970*, (Arlington, VA.: ERIC Document Reproduction Service, ED 059 415, 1972.)

3. D. Hawkridge, op cit., 1968, pp. 30-31.

4. For a concise review of the problems associated with measuring achievement in grade equivalencies see D. Horst's "What's Bad About Grade Equivalent Scores," Mountain View, California: RMC Research Corporation, October, 1976.

5. There appears to be less controversy over the 2:3 ratio per se than there is over how much of the achievement deficit may occur during the summer.

6. Wargo, 1971, op cit., p. 5.

7. J. McV. Hunt, *Intelligence and Experience*, New York: Ronald Press, 1961.

   . B. Bloom, *Stability and Change in Human Characteristics*, New York: John Wiley and Sons, Inc., 1964.

8. B. Bloom, op cit., pp. 33-34.
   J. McV. Hunt, op cit., pp. 18-21.

9. See L. Erlenmeyer-Kimling and L. Jarvik, "Genetics and Intelligence: A Review," *Science*, 142, 1963, pp. 1477-79, for a summary of many of the studies of Hunt and Bloom and a graphic analysis showing I.Q. variations of unrelated persons, foster parents and children, natural parents and children, siblings and twins.

10. H. Newman, F. Freeman and K. Holzinger, *Twins, A Study of Heredity and Environment*, Chicago: The University of Chicago Press, 1937.

11. See a charting of the twins in the Chicago Group Study by A. Anatasi, in *Differential Psychology*, New York: MacMillan, 1956, p. 299.

12. L. Erlenmeyer and L. Jarvik, op cit., pp. 1477-79.

13. See a review of this research by A. Jensen in "The Heritability of Intelligence," Readings in Child Development, Second Edition, H. Munsinger (Ed.), New York: Holt, Rinehart and Winston, 1975, pp. 131-135.

14. Ibid.

15. A. Anatasi, op cit., p. 299.

16. H. McV. Hunt, op cit., pp. 19-21.

17. R. Spitz, "Hospitalism, An Inquiry into the Genesis of Psychiatric Conditions in Early Childhood," Psychoanalytic Study of the Child, v1, 1945, pp. 53-74.

18. Ibid.

19. H. Skeels and H. Dye, "A Study of the Effects of Differential Stimulation on Mentally Retarded Children," Proceedings of the American Association of mental Deficiency, v44, 1939, pp. 114-136.

20. H. Skeels, "Effects of Adoption on Children from Institutions," Children, January-February, 1965, pp.33-34.

H. Skeels, "Adult Status of Children with Contrasting Early Life Experiences: A Follow-Up Study," Child Development Monographs, v31, n3, Serial No. 105, 1966.

21. H. Skeels, op cit., p. 34.

22. J. Anderson, "The Limitations of Infant and Pre-School Tests in the Measurement of Intelligence," Journal of Psychology, v8, 1939, pp. 851-79.

N. Bayley, "Consistency and Variability in the Growth of Intelligence from Birth to Eighteen Years," Journal of Genetic Psychology, v75, 1949, pp. 165-96.

E. Ebert and K. Simmons, "The Brush Foundation Study of Child Growth and Development, I. Psychometric Tests," Monographs of the Society for Research in Child Development, v8, n21, 1943.

F. Freeman and C. Flory, "Growth in Intellectual Ability as Measured by Repeated Tests," Monographs of the Society for Research in Child Development, v2, n2, 1937.

M. Honzik, J. Macfarlane and L. Allen, "The Stability of Mental Test Performance between Two and Eighteen Years," Journal of Experimental Education, v17, 1938, pp. 309-24.

L. Sontag, C. Baker and V. Nelson, "Mental Growth and Personality: A Longitudinal Study," Monographs of the Society for Research in Child Development, v23, n2, 1958.

23. B. Bloom, op cit., pp. 56-57.

24. Ibid., p. 57.

25. J. McV. Hunt, op cit., pp. 24-25.

26. O. Klineberg, Negro Intelligence and Selective Migration, New York: Columbia University Press, 1935.

27. E. Lee, "Negro Intelligence and Selective Migration: A Philadelphia Test of the Klineberg Hypothesis," American Sociological Review, v16, 1951, pp. 227-33.

28. L. Wheeler, "A Comparative Study of the Intelligence of East Tennessee Mountain Children," Journal of Educational Psychology, v33, 1942, pp. 321-34.

29. O. Lewis, La Vida. A Puerto Rican Family in the Culture of Poverty, San Juan and New York, New York: Random House, 1965, p. ii.

30. O. Lewis, The Children of Sanchez: Autobiography of a Mexican Family, New York: Random House, 1961, p.xxx.

31. F. Riessman, The Culturally Deprived Child, New York: Harper, 1962.

32. J. Conant, Slums and Suburbs, New York: McGraw-Hill, 1961.

33. K. Clark, Dark Ghetto, New York: Harper & Row, 1965.

34. See, for example, B. Bloom, A. Davis and R. Hess's accounting of the subject in Compensatory Education for Cultural Deprivation, New York: Holt, Rinehart and Winston, Inc., 1965, pp. 30-31.

35. D. Ausubel and P. Ausubel, "Ego Development Among Segregated Negro Children," Education in Depressed Areas, A. Harry Passow (Ed.), New York: Teachers College Press, Columbia University, 1963.

36. Ibid., p. 174.

37. M. Rosenberg, Society and the Adolescent Self Image, Princeton, New Jersey: Princeton University Press, 1965, p. 65.

38. F. Boyd, "The Levels of Aspirations of White and Negro Children in a Non-Segregated Elementary School," Journal of Social Psychology, v36, first half, 1952, pp. 191-96.

R. Sears, Patterns of Child Rearing, Evanston: Row and Peterson, 1952.

M. Weinger and W. Murray, "Another Lock at the Culturally Deprived and Their Levels of Aspiration," Journal of Educational Sociology, v36, 1963, pp. 319-21.

39. Many have associated Siegfried Engelman with this position because of his insistance that disadvantaged children need language training programs which will effectively teach ". . . language and reasoning skills." For an accounting of how to teach language and reasoning see S. Engleman's "How to Construct Effective Language Programs for the Poverty Child," in Language and Poverty, F. Williams, (Ed.), Chicago: Markham Publishing Company, 1970, pp. 102-121.

40. B. Bernstein, "Linguistic Codes, Hesitation, Phenomena and Intelligence," Language and Speech, v5, Part 1, 1962, pp. 31-46.

41. D. Augusbel, "How Reversible are the Cognitive and Motivational Effects of Cultural Deprivation? Implications for Teaching the Culturally Deprived Child," Urban Education, vl, Summer, 1964, pp. 16-38.

42. V. John, "Intellectual Development of Slum Children: Some Preliminary Findings," American Journal of Orthopsychiatry, v33, 1963, pp. 813-22.

43. For a more complete description of the rationale for and implementation of a language training program see C. Bereiter and S. Englemann's Teaching Disadvantaged Children in the Preschool, Englewood Cliffs, New Jersey: Prentice Hall, 1966.

44. M. Deutsch, "Facilitating Development in the Pre-School Child: Social and Psychological Perspectives," Merrill-Palmer Quarterly, v10, 1964, pp. 249-63.

45. D. Hebb, "The Motivating Effects of Extrareceptive Stimulation," American Psychologist, v13, 1958, p. 109.

46. M. Von Senden, Raum and Gestaltauffassung bei operierten Blindegeborener von und nach der Operation Leipzig: Barth, 1932, cited by Cynthia P. Deutsch in "Environment and Perceptions," in Social Class, Race, and Psychological Development, M. Deutsch, I. Katz and A. Jensen (Eds.), New York: Holt, Rinehart and Winston, 1968, pp. 71-72.

47. B. White, "Informational Education during the first Months of Life," paper presented at the Social Science Research Council Conference on Pre-School Education, Februry 8, 1966, cited by Cynthia Deutsch in "Environment and Perception," in Social Class, Race and Psychological Development, M. Deutsch, I. Katz and A. Jensen (Eds.), New York: Holt, Rinehart and Winston, 1968, pp. 72-73.

48. M. Deutsch and P. Katz, "The Disadvantaged Child and the Learning Process," Education in Depressed Areas, A. Harry Passow (Ed.), New York: Teachers College Press, Columbia University, 1963, pp. 163-79.

49. Ibid., pp. 170-171.

50. C. Deutsch, "Auditory Discrimination and Learning: Social Factors," Merrill-Palmer Quarterly, v10, 1964, pp. 277-95.

51. M. Deutsch and P. Katz, op cit., p. 163.

52. B. Bloom, op cit, p. vii.

53. Ibid., p. 72.

54. Ibid.

55. Ibid.

56. J. McV. Hunt, "Environment, Development and Scholastic Achievement," a paper prepared for the Arden Conference on Pre-School Enrichment of Socially Disadvantaged Children in December, 1963, reprinted in Social Class, Race and Psychological Development, M. Deutsch, I. Katz and A. Jensen (Eds.), New York: Holt, Rinehart and Winston, 1968, pp. 293-296.

57. Ibid.

58. A. Jensen, "Learning in the Pre-School Years," The Journal of Nursery Education, v18, n2, 1963, p. 133.

59. Ibid., pp. 133-139.

60. E. Lee, op cit., pp. 227-233.

61. J. Kirk, Early Education of the Mentally Retarded, Urbana, Illinois: University of Illinois Press, 1958, p. 209.

62. Ibid., pp. 209-211.

63. J. Bowlby, Maternal Care and Mental Health, Geneva: World Health Organization, 1952.

64. S. Province and R. Lipton, Infants in Institutions, New York: International Universities Press, 1962, p. 144.

65. W. Dennis, "Causes of Retardation Among Institutional Children: Iran," The Journal of Genetic Psychology, v96, first half, 1960, pp. 56-57.

66. W. Goldfarb, "The Effects of Early Institutional Care on Adolescent Personality," The Journal of Experimental Education, v12, p. 143(A).

67. H. Mann, The 12th Annual Report to the Massachusetts Board of Education, 1848, reprinted in Foundations of Education in America, J.W. Noll and S. P. Kelly (Eds.), New York: Harper and Row, 1970, pp. 120-15.

68. C. Jencks, et al., Inequality, New York: Harper and Row, 1972, p. 87.

69. Ibid.

70. A. Passow, (Ed.), Education in Depressed Areas, New York: Teachers College Press, Columbia University, 1963, p. vii.

71. Ibid., p. 351.

72. Ibid., p. 18.

73. Ibid., p. 89.

74. Ibid., p. 144-145.

75. See B. Bloom, A. Davis and R. Hess, Compensatory Education for Cultural Deprivation, New York: Holt, Rinehart and Winston, 1965.

76. Ibid., p. 4.

77. Ibid., p. 6.

78. M. Deutsch, op cit., 1964, p. 260.

79. Ibid., pp. 249-263.

80. K. Clark, "Educational Stimulation of Racially Disadvantaged Children," Education in Depressed Areas, A. Harry Passow (Ed.), New York: Teachers College Press, Columbia University, 1963, p. 157.

81. Ibid.

82. Great Cities School Improvement Studies, Ford Foundation Project, mimeographed, 1960, reprinted in The Culturally Deprived Child by F. Riessman, New York: Harper and Row, 1962, pp. 125-6.

83. P. Sexton, Education and Income, New York: Viking, 1961.

84. F. Reissman, op cit.

85. R. Freeman, "The Alchemists in Our Public Schools," quoted by J. Ashbrook in the House of Representatives, Congressional Record, April 24, 1969, pp. 10325-26.

86. Ibid.

87. United States Commission on Civil Rights, 1967, Racial Isolation in the Public Schools, V.I, U.S. Government Printing Office, Washington, D.C., February, 1967.

88. H. Passow, op cit., p. 343; R. Freeman, op cit., p. 10326.

89. Ibid.

90. U.S. Commission on Civil Rights, op cit., pp. 124-125.

91. Ibid., p. 123.

92. Ibid.

93. E. Gordon and D. Wilkerson, Compensatory Education for the Disadvantaged, New York: College Entrance Examination Board, 1966, p. 250.

94. Baynham, 1963, quoted in B. Stickney, Compensatory Education and Pupil Achievement, Unpublished Doctoral Dissertation, University of Massachusetts, Amherst, 1976, p. 110.

95. M. Almy, The Early Childhood Educator at Work, New York: McGraw-Hill, Inc., 1975, p. 84.

96. Stickney, op cit., pp. 186-193.

97. Ibid., pp. 188-190.

98. Almy, op cit., p. 88.

99. K. Hecht, "Title I Federal Evaluation: The First Five Years," Teachers College Record, v75, nl, 1973, pp. 67-78.

100. Ibid.

101. Ibid., p. 67.

102. See Wargo, op cit., 1971, p.5.

103. See A. Anatasi, op cit., p. 299, cited in B. Bloom, op cit., 1964, p. 70.

104. Webster's New International Dictionary of the English Language, Second Edition, Springfield, Massachusetts: Unabridged, G. & C. Merriam Company, Publishers, 1951, p. 545.

105. J. Coleman, et al., Equality of Educational Opportunity, Washington, D.C.: U.S. Department of Health, Education and Welfare, U.S. Government Printing Office, 1966.

106. Ibid., p. iii.

107. F. Mosteller and D. Moynihan, (Eds.), On Equality of Educational Opportunity, New York: Random House, 1972.

108. J. Coleman, op cit.

109. Ibid.

110. Ibid.

111. Ibid., p. 122.

113. Ibid., p. 113.

114. Ibid., p. 274.

115. Ibid., pp. 317-318.

116. Ibid., p. 23.

117. Ibid., p. 32.

118. Ibid., p. 22.

119. Ibid., pp. 298-302.

120. Ibid., p. 300.

121. Ibid., p. 325.

122. K. Hecht, op cit.

123. See F. Mosteller and D. Moynihan, op cit., for a reanalyses generally supportive of Coleman's conclusions.

124. See S. Bowles and H. Levin, "The Determinants of Scholastic Achievement--An Apprisal of Some Recent Evidence," The Journal of Human Resources, v8, Winter, 1968, pp. 3-24.

E. Hanushek and J. Kain, "On the Value of Equality of Educational Opportunity as a Guide to Public Policy," On Equality of Educational Opportunity, F. Mosteller and D. Moynihan (Eds.) New York: Random House, 1972, pp. 116-145.

BIBLIOGRAPHY

Part II

1. J. Coleman, et al., Equality of Educational Opportunity, Washington, D.C.: U.S. Government Printing Office, 1961, p. 22.

2. See the introduction to D. Moynihan and F. Mosteller's, On Equality of Educational Opportunity Survey, New York: Vintage Books, 1972.

3. M. McLaughlin, Evaluation and Reform: The Elementary and Secondary Education Act of 1965, Title I, San Monica, California: The Rand Corporation, 1974, p. 3.

4. Ibid., p. 16.

5. Ibid.

6. Ibid., Congressional Liaison from p. 16; ". . . provisions for appropriate objectives . . . " taken from Public Law No. 89-10, Sec. 205, (The Elementary and Secondary Education Act), passed by Congress April 11, 1965 and quoted in McLaughlin, op cit., p. 17.

7. Ibid.

8. Ibid.

9. The States Report--The First Year of Title I, Elementary and Secondary Education Act of 1965, (Arlington, VA.: ERIC Document Reproduction Service, ED 021 946, 1967).

10. The Second Annual Report of Title I of the Elementary and Secondary Education Act of 1965, School Year, 1966-67, (Arlington, VA.: ERIC Document Reproduction Service, ED 021 946, 1967).

11. From a personal interview with Kathryn Hecht in Denver, Colorado, June 5, 1980.

12. Ibid.

13. Education of the Disadvantaged: An Evaluative Report on Title I, Elementary and Secondary Education Act of 1965, Fiscal Year 1968, (Arlington, VA.: ERIC Document Reproduction Service, ED 047 033, 1970).

14. Ibid.

15. Ibid., pp. 97 & 375.

16. M. McLaughlin, op cit., p. 57.

17. Ibid., p. 59.

18. The Impact of Head Start: An Evaluation of the Effects of Head Start on Children's Cognitive and Affective Development, (Arlington, VA.: ERIC Document Reproduction Service, ED 036 321, 1969).

19. Quoted by John Beckler in "Congress and the Administration Scrutinize the Results of Title I," School Management, v14, n9, 1970, p. 6.

20. See D. Hawkridge, et al., A Study of Exemplary Programs for the Education of Disadvantaged Children, Part I, (Arlington, VA.: ERIC Document Reproduction Service, ED 023 776, 1968, pp. 1-2).

    K. Tallmadge, The Development of Project Information Package for Effective Approaches in Compensatory Education, Final Report, (Arlington, VA.: ERIC Document Reproduction Service, ED 099 457, 1974, pp. ii, v, vi).

21. United States Commission on Civil Rights, 1967, Racial Isolation in the Public Schools, Vol. I., Washington, D.C., U.S. Government Printing Office, February, 1967, pp. 121, 125, 127.

22. In 1967 the American Institute for Research (AIR) was in the process of searching for exemplary programs, under contract with the U.S. Office of Education. In the first AIR search over 1,000 projects were reviewed, a second search (published in 1969) extensively combed the literature (mailing letters of inquiry to 320 programs) and a third search (published in 1971) reviewed over 1,200 programs.

    See D. Hawkridge, et al., op cit., p. 1; D. Hawkridge, et al., A Study of Further Exemplary Programs for the Education of Disadvantaged Children, Part II, (Arlington, VA.: ERIC Document Reproduction Service, ED 036 668, 1969, p. 6); M. Wargo, et al., Further Examination of Exemplary Programs for Educating Disadvantaged Children, (Arlington, VA.: ERIC Document Reproduction Service, ED 055 128, 1971).

23. It Works Series: Summaries of Selected Compensatory Education Projects, Washington, D.C.: U.S. Government Printing Office, 1970.

24. D. Hawkridge, et al., op cit., 1968, p.1.

25. D. Hawkridge, et al., op cit., 1969, p. 6.

26. 1:1 Gains

27. M. Wargo, et al., ESEA Title I: A Reanalysis and Synthesis of Evaluation Data from Fiscal Year 1965 to 1970, (Arlington, VA.: ERIC Document Reproduction Service, ED 059 415, 1972, p. 180).

28. M. Wargo, op cit., p. iii.

29. Ibid., pp. iii-iv.

30. M. Wargo, 1971, op cit., p. 44.

31. Ibid., pp. 185-186.

32. M. McLaughlin, et al., The Effects of Title I ESEA: An Exploratory Study, (Arlington, VA.: ERIC Document Reproduction Service, ED 073 216, 1971, p. 4.

33. Ibid., p. 2.

34. Ibid., pp. 1-8 and 13.

35. Ibid., p. 17.

39. Ibid., p. 18.

40. Ibid., pp. 13-14.

41. Ibid., p. 16.

42. Ibid., p. 13.

43. C. Foat, Selecting Exemplary Compensatory Education Programs for Dissemination Via Project Information Package, (Arlington, VA.: ERIC Document Reproduction Service, ED 099 456, 1964, p. 14).

44. Ibid., pp. 14, 31-52.

45. Ibid., pp. 31-52.

46. From a telephone interview with Ann Weinheimer, Evaluation Specialist in the Office of Evaluation, Department of Education, June 19, 1980.

47. Ibid.

48. Ibid.

49. Ibid.

50. Ibid.

51. Educational Programs that Work, San Francisco, California: Far West Laboratory, Fall, 1980.

52. Ibid.

53. M. Stearns, et al., Evaluation of the Field Test of Project Information Packages, Volume I, (Arlington, VA.: ERIC Document Reproduction Service, ED 142 568, 1977, p. 36.)

54. Weinheimer, op cit.

55. S. Ryan (Ed.), A Report on Longitudinal Evalutions of Pre-School Programs, Vol. 1: Longitudinal Evaluations, (Arlington, VA.: ERIC Document Reproduction Service, ED 093 500, 1974.)

56. See C. Bereiter and S. Engelmann, Teaching Disadvantaged Children in the Preschool, Englewood Cliffs, New Jersey: Prentice-Hall, 1966; D. Weikart, "Relationship of Curriculum, Teaching and Learning in Pre-School Evaluation," Preschool Programs for the Disadvantaged, J.C. Stanley (Ed.), Baltimore, Maryland: The Johns Hopkins University Press, 1972, pp. 22-66.

57. See D. Weikart, *et al.*, *Longitudinal Results of the Ypsilanti Perry Preschool Project, Final Report*, (Arlington, VA.: ERIC Document Reproduction Service, ED 044 536, 1970); D. Weikart, 1972, op cit., pp. 22-26.

58. S. Gray and R. Klaus, "The Early Training Project: A Seventh-Year Report," *Child Development*, v41, 1970, pp. 909-24.

59. E. Herzog, C. Newcomb and I. Cisin, "Double Deprivation: The Less They Have, the Less They Learn," in *A Report on the Longitudinal Evaluations of Preschool Programs, Vol. I: Longitudinal Evaluations*, by Sally Ryan, (Arlington, VA.: ERIC Document Reproduction Service, ED 093 500, pp. 53-69.)

60. M. Karnes, *et al.*, "The Karnes Pre-School Program: Rationale, Curricular Offerings and Follow Up," in *A Report on Longitudinal Evaluations of Pre-School Programs, Vol. I: Longitudinal Evaluations*, by Sally Ryan, (Arlington, VA.: ERIC Document Reproduction Service, ED 093 500, 1974, pp. 95-108.)

61. E. Beller, "The Impact of Early Education on Disadvantaged Children," in *A Report on Longitudinal Evaluations of Pre-School Programs, Vol. I: Longitudinal Evaluations*, by Sally Ryan, (Arlington, VA.: ERIC Document Reproduction Service, ED 093 500, 1974, pp. 15-48.)

62. See V. Van De Riet and M. Resnick, *Developmental and Educational Implications of a Successful Model of Compensatory Education, The Learning to Learn Program*, (Arlington, VA.: ERIC Document Reproduction Service, ED 082 852. 1973); H. Springle, "Learning to Learn Program" in *A Report on Evaluations of Preschool Programs, Vol. I: Longitudinal Evaluations*, by Sally Ryan, (Arlington, VA.: ERIC Document Reproduction Service, ED 093 500, 1974, pp. 109-124.)

63. M. Deutsch, *et al.*, a "Brief Synopsis of an Initial Enrichment Program in Early Childhood," in *A Report on Evaluations of Preschool Programs, Vol. I: Longitudinal Evaluations*, by Sally Ryan, (Arlington, VA.: ERIC Document Reproduction Service, ED 093 500, 1974, pp. 49-60.)

64. Ibid., p. 15.

65. Ibid.

66. Wargo, op cit., p. 44.

67. R. North, *et al.*, *Evaluation Report for the Project 'More Effective Schools Program' in Poverty Area Schools*, 1968-69, (Arlington, VA.: ERIC Document Reproduction Service, ED 041 998, 1969, pp. 109-110.)

68. R. Freeman, "The Alchemists in our Public Schools," quoted by J. Ashbrook in the House of Representatives, *Congressional Record*, April 24, 1969, p. 10326.)

69. T. Thomas & S. Delavin, *Patterns in ESEA Reading Achievement*, (Arlington, VA.: ERIC Document Reproduction Service, ED 120 087, March, p. 39.

70. Ibid., pp. 38-41.

71. Ibid., pp. 37-38.

72. Ibid., pp. 40-41.

73. Ibid., pp. 42-43.

74. Ibid., p. 43.

75. See L. Stebbins, et al., Education As Experimentation: A Planned Variation Model, Volume IV-A, An Evaluation of Follow Through, Cambridge, Massachusetts: Abt Associates Inc.

76. See W. Abelson and E. Zigler, "Effects of a Four-Year Follow-Through Program on Economically Disadvantaged Children," Journal of Educational Psychology, v66, n5, pp. 756-771.

77. G. Hendrickson, "Review of Title I Evaluation Studies," paper presented at the annual meeting of the American Educational Research Association, New York, New York, April 10-15, 1977, p. 38.

78. Weinheimer, op cit.

79. G. Tallmadge, "Interpreting NCEs," Mountain View, California: RMC Research Corporation, October, 1976, p.1.

80. D. Horst, "What's Bad about Grade Equivalent Scores," Mountain View, California: RMC Research Corporation, October, 1976, p. 2.

81. Ibid., pp. 2-4.

82. Ibid., p.3.

83. Tallmadge, op cit., p. 3.

84. Ibid., pp. 2-4.

85. Ibid., p. 6.

86. The authors learned of the RMC "working paper" during a telephone conversation (June, 19, 1980) with OE's Janice Anderson, an Evaluation Specialist, who has been closely affiliated with Title I evaluation. After learning of our interest in the new Title I reporting system, Ms. Anderson graciously mailed us a copy of the paper.

87. G. Tallmadge and B. Fagan, "Cognitive Growth and Growth Expectations in Reading and Mathematics," Mountain View, California: RMC Research Corporation, prepared for U.S. DHEW's OPOE, November 1977, p. 4.

88. G. Mayeske, A Study of Compensatory Reading Programs: Technical Summary, Washington, D.C.: USOE, OPBE, Elementary and Secondary Programs Division, September, 1976, p. 61, quoted in Robert Rossi, et al.'s Summaries of Major Title I Evaluations, (Arlington, VA.: ERIC Document Reproduction Service, ED 145 012, p. 128.)

89. From a telephone conversation with George Mayeske, June 25, 1980.

90. Ibid.

91. P. Hill and I. Rotberg, The Effects of Services on Student Development, Washington, D.C.: U.S. Government Printing Office, September 30, 1977.

92. Ibid, pp. 19-20.

93. Compensatory Education Study, Washington, D.C.: U.S. Government Printing Office, September, 1978, p. 83.

94. Ibid., p.84.

95. Ibid., pp.84 & 87.

96. M. Wang, "Evaluating the Effectiveness of Compensatory Education," paper presented at the annual meeting of the American Educational Research Corporation, Boston, Massachusetts, April 10, 1980.

97. Ibid.

98. L. Carter, "The Sustaining Effects Study: Evaluating the Effectiveness of Compensatory Education," a presentation at the Second Colorado Title I Conference, Colorado Springs, Colorado, January 4, 1980.

99. R. Darlington, et al., "Preschool Programs and Later School Competencies," Science, v208, n4440, pp. 202-204.

99a. P. DeVito and J. Rubenstein, "A Follow-Up Study of Title I Participants in Rhode Island," paper presented at the annual meeting of the American Educational Research Association, San Francisco, California, April 8-13, 1979, p. 5.

99b. Ibid., p. 8.

100. S. Powers, "A Comparison of Cognitive Gains of Seventh, Eighth and Ninth Grade Students Across Two Years of Title I," paper presented at the annual meeting of the American Educational Research Association, Boston, Massachusetts, April 7-11, 1980, p. 14.

101. A. Weinheimer, op cit.

102. Ibid.

103. "British Researchers look at Urban Secondary Schools and their Effects on Disadvantaged Children," Education Times, Monday, January 28, 1980, p. 2.

104. From a telephone interview with Ida Brooks of the National Association for Educational Progress, June 26, 1980.

105. R. Edmunds, "Effective Schools for the Urban Poor," Educational Leadership, v37, n1, October, 1978, p. 22.

106. G. Mayeske, op cit.

107. S. Powers, op cit., p. 14.

108. B. Rosenshine, "Academic Engaged Time, Content Covered, and Direct Instruction," Journal of Education, v160, n3, August, 1978, p. 46.

109. See R. Burns' "Mastery Learning: Does It Work," Educational Leadership, v37, n2, 1979, pp. 110-113.

110. Ibid.

111. Ibid.

112. W. Aikin, The Story of the Eight Year Study, New York: Harper & Brothers, 1942.

113. P. Perry, "Characteristics of Successful Title I Programs," from a presentation at the Second Colorado Title I Conference, Colorado Springs, Colorado, January 4, 1980.

114. From a telephone interview with Chuck Sekera, June 27, 1980.

115. Ibid.

116. From a telephone interview with Peggy Seavy, June 27, 1980.

117. Ibid.

118. C. Sekera, op cit.

119. From an interview with Larry Bussey, June 16, 1980.

120. From a telephone interview with OPOE's Janice Anderson, September, 1980.